COMPARISON OF ENSEMBLE MEAN AND DETERMINISTIC FORECASTS FOR LONG-RANGE AIRLIFT FUEL PLANNING

THESIS

Haley A. Homan, Captain, USAF

AFIT-ENP-14-M-15

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-14-M-15

COMPARISON OF ENSEMBLE MEAN AND DETERMINISTIC FORECASTS FOR LONG-RANGE AIRLIFT FUEL PLANNING

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science

Haley A. Homan, BS

Captain, USAF

27 Mar 2014

AFIT-ENP-14-M-15

COMPARISON OF ENSEMBLE MEAN AND DETERMINISTIC FORECASTS FOR LONG-RANGE AIRLIFT FUEL PLANNING

Haley A. Homan, BS
Captain, USAF

Approved:

//signed//	17 March 2014
Lt Col Robert S. Wacker, PhD (Chairman)	Date
//signed//	17 March 2014
Lt Col Kevin S. Bartlett, PhD (Member)	Date
//signed//	17 March 2014
Ariel O. Acebal, PhD (Member)	Date

Abstract

Implementing an ensemble mean forecast to aid in fuel planning for long-range strategic airlift has the potential to improve upon the deterministic forecasts currently used. More accurate wind forecasts could aid in a significant reduction in annual fuel costs for the DoD. This study focuses on the wind forecasts from the Global Forecast System (GFS) deterministic model and the ensemble mean wind forecasts from the Global Ensemble Forecast System (GEFS), Global Ensemble Prediction System (GEPS), and Mesoscale Ensemble Prediction System (MEPS) over a 60-day period from 19 Sep through 17 Nov 2013. The fuel burn and total spread were computed for five great circle flight routes and five aircraft using each model's wind data. The deterministic fuel burn error was then compared to the ensemble mean fuel burn error. For each of the routes of flight investigated at cruise levels 500mb (FL180) and 250mb (FL340), the amount of reserve fuel required to account for the uncertainty in the wind forecasts was typically lower for the ensemble mean forecasts during forecast hours 12 to 48.

Acknowledgements

First and foremost I would like to thank my mom, dad, sister, and brother for their love and support. Your words of encouragement over the past 21 months truly helped get me through this challenging program. Next, I want to thank my advisor, Lt Col Wacker, as well as my other committee members, Lt Col Bartlett, and Dr. Acebal for all of their help and support not only during my thesis research, but throughout my entire time here at AFIT. Many thanks go out to Evan Kuchera from AFWA's 16th Weather Squadron for not only providing me with the AFWA ensemble data needed for this study, but also for answering all of the questions I had regarding ensembles. Last, but not least, I'd like to thank my classmates, Jeremy Hromsco, Brad Clements, Kyle Thurmond, and Coy Fischer. I couldn't think of a better group of people to share the "AFIT experience" with than you guys. It was great to have such close friends to spend countless hours with doing homework and studying for tests. Oh, and who could forget the labs during our final quarter? You know, the ones that only took 20 minutes to complete! Thank you for the memories and best of luck to you at your next assignment and whatever else life brings your way. God Bless!

Haley A. Homan

Contents

List of Figures

Figure Page

List of Tables

COMPARISON OF ENSEMBLE MEAN AND DETERMINISTIC FORECASTS FOR LONG-RANGE AIRLIFT FUEL PLANNING

1. Introduction

Accurate wind forecasts are essential to fuel planning for Air Mobility Command (AMC). Poor wind forecasts can result in over- or under-estimating required fuel loads, which translates into wasted money and man hours for the Air Force. Currently, wind-optimized flight planning for AMC is performed using Advanced Computer Flight Planning (ACFP) system. ACFP incorporates wind and temperature data from numerical weather prediction (NWP) models. To date, only deterministic forecast data has been used.

Until recently, the convention for meteorologists has been to use NWP models to produce a deterministic forecast. A deterministic forecast is one in which a single model generates an estimate of the current atmospheric state at a particular time (initial conditions) and then integrates the governing equations (conservation of mass, conservation of momentum, conservation of energy, a conservation equation for water mass, and the equation of state) over time in order to produce a single solution that represents the future state of the atmosphere. Despite improved physics packages, more computing power, finer resolution, and improved numerical methods, there is still a considerable amount of uncertainty associated with deterministic forecasts (Leutbecher and Palmer 2008). Because of this uncertainty, the use of NWP deterministic forecasts in optimized flight planning systems may not be the most effective approach.

An ensemble of models enables users to quantify the amount of uncertainty in

NWP forecasts and visualize it through various graphical products. Each member of an ensemble prediction system (EPS) is created by either perturbing the initial conditions or employing varying physics parameterizations within the model itself. This research is focused on determining if the mean wind forecast from an ensemble of models would be more advantageous than a single model deterministic forecast when determining flight routes .

Ensemble forecast uncertainty can be presented in a number of ways. Some of these include: (1) spaghetti diagrams, which give a human forecaster an idea of the probability distribution for the forecast, (2) probabilities of exceeding specific thresholds, such as wind gusts greater than 35 knots, and (3) forecast mean and spread. The mean is simply the average value of all the ensemble members' forecasts of a parameter, while the spread quantifies the uncertainty via the standard deviation among the ensemble members' forecasts. Several authors have shown that the ensemble mean forecast performs better, on average, than a deterministic forecast (Anderson 1996; Eckel et al. 2008; Toth and Kalnay 1993; Tracton and Kalnay 1993). Keith and Leyton (2007) showed that by switching from the traditional terminal aerodrome forecast, which is a deterministic forecast, to a probabilistic forecast for adverse landing weather conditions, commercial airlines could see significant savings in fuel expenses. Additionally, several other studies have shown the value of using ensemble-based probabilistic decision inputs over deterministic or climatological information (Katz and Murphy 1997; Richardson 2000; Palmer 2002; Zhu et al. 2002).

The main objectives of this research are to: (1) determine if an ensemble mean wind forecast is more advantageous than a deterministic forecast for strategic airlift fuel planning and (2) determine if a correlation exists between the ensemble spread and the error in calculated fuel burns. The remainder of this document is

organized into four chapters. Chapter two provides background information and begins with a brief history of numerical weather prediction. Next it defines what an ensemble is and the processes involved in creating an ensemble. Chapter three discusses the methods used to obtain and analyze the data used in this study. Chapter four presents the results. Finally, chapter five summarizes the study and provides recommendations for future work.

2. Background

2.1 Ensemble Prediction System (EPS)

2.1.1 History of NWP.

NWP has a rich history that can be traced back to 1949, when Jule Charney and his group produced the first successful 24-hr forecast (Lewis 2005). During this time, NWP principles were based on determinism, which simply stated that the future state of a system is completely determined by the initial state of the system (Lewis 2005). It was the success of Charney's group that sparked the meteorological community across the world to begin working on improving the NWP framework in hopes of being able to produce extended-range operational weather forecasts.

During the mid-1950s and throughout the 1960s, primitive equation (PE) models took the place of quasi-geotrophic models as the basis for operational short-range NWP and were further explored by Charney (1955) and Richardson (1965). The quasi-geostrophic approximation has the effect of filtering out sound and gravity waves (Kalnay 2003). It was this approximation which made the first successful 24-hour forecast by Charney possible. Although the use of PEs had distinct advantages over the quasi-geostrophic models, the computational requirements needed to account for the finer time steps and increased number of variables made it impossible to implement operationally until 1965 (Lewis 2005). Leith was the first to develop a PE-based global circulation model (GCM) in the summer of 1960. The model extended from the equator to 60° N and incorporated a 5° latitude x 5° longitude horizontal grid over five vertical levels. It was also the first numerical model to include moisture, clouds, and rain (Leith et al. 1965). During this time period, Lorenz made a monumental contribution to NWP development. Beginning in 1956,

he set out to disprove the hypothesis that a linear regression could be used to adequately forecast under the constraint of non-linear dynamics. He ultimately discovered a non-periodic evolution of the atmosphere commonly known today as chaos theory, which is fully described in his scientific biography *The Essence of Chaos* (Lorenz 1993). As a result of Lorenz's findings, by 1964, meteorologists began shifting their focus toward predictability and the need for statistical consideration in order to achieve useful long-range forecasts (Lewis 2005). It was also during this time that Lorenz shared his vision of ensemble forecasting with the meteorological community:

> The proposed procedure chooses a finite ensemble of initial states, rather than a single observed initial sate. Each state within the ensemble resembles the observed state closely enough so that the differences might be ascribed to errors or inadequacies in observation. A system of dynamic equations previously deemed to be suitable for forecasting is then applied to each member of the ensemble, leading to an ensemble of states at any future time. From an ensemble of future states, the probability of occurrence of any event, or such statistics as the ensemble mean and ensemble standard deviation of any quantity may be evaluated. Between the near future, when all states within an ensemble will look about alike, and the very distant future when the two states within an ensemble will show no more resemblance than two atmospheric states chosen at random, it is hoped that there will be an extended range when most of the states in an ensemble while not constituting good pin-point forecasts, will possess certain important features in common. It is for this extended range that the procedure may prove useful. (Lorenz 1965).

The next big contribution to probabilistic NWP and forecast uncertainty came from Epstein (1969) and his work on a stochastic-dynamic (SD) approach to NWP. SD prediction was conducted using Monte Carlo methods, which assume a known randomly sampled probability distribution function (PDF). For much of the time from the 1960s to the mid-1980s meteorologists were waiting for the operational implementation of SD forecasting, which relied on the availability of affordable

parallel-processing computers. Further study occurred on the limitations of predictability, computing issues related to the Monte Carlo and SD approximations, the perturbation methodology, and finally, flow-dependent uncertainty (Lewis 2005). It was during this time period that Leith proposed using multiple NWP runs to produce a limited sample of future atmospheric states. Leith (1974) was able to show improvements in forecasts with lead times out to 10 days by using the mean values from approximately 10 different NWP runs.

During the early 1990s, the European Center for Medium-Range Weather Forecasts (ECMWF), National Centers for Environmental Prediction (NCEP), and the Meteorological Service of Canada (MSC) first put operational stochastic-dynamic prediction, (i.e. ensemble forecasts) into operational use. An initial comparison of these three EPSs can be found in Buizza et al. (2005), however, for a more comprehensive list of updates and the most up to date information regarding each EPS, refer to their respective websites (`www.ecmwf.int`, `www.ncep.noaa.gov`, and `http://weatheroffice.ec.gc.ca/ensemble/index_e.html`).

Since the operational implementation of EPSs, advances in computing power have expanded the number of ensemble members and increased the vertical and horizontal resolution of the members. Improvements to model physics and data assimilation schemes have also reduced initial state errors.

2.1.2 Definition of EPS.

Eckel et al. (2008) explain that the purpose of an ensemble forecast is to simulate error growth associated with errors in the analysis of the initial state, and the deficiencies in the NWP model to produce a sample of likely forecast states. EPSs accomplish this by producing multiple forecasts through a number of different methods. One method used to create multiple forecasts is to perturb the sur-

face or lateral boundary conditions (note that lateral boundary conditions can only be perturbed in regional models). Another method involves perturbing the initial conditions. A final method involves perturbing the model itself. The model can be perturbed by changing the dynamical formulation by, for example, using different vertical coordinate types. Another way to perturb the model is to change the numerical methods it uses. An example of this would be to use grid-point differencing versus spectral differencing. A third method to perturb the model is to change one or more of the physical parameterizations, which are used to represent atmospheric phenomenon that cannot be explicitly resolved by the model. An example of such a parameterization is the cumulus convective scheme. A final way to perturb the model is to simply change the horizontal and/or vertical resolution (UCAR 2004).

The three different EPSs used in this study include NCEP's Global Ensemble Forecast System (GEFS), and the Air Force Weather Agency's (AFWA) Global Ensemble Prediction System (GEPS) and Mesoscale Ensemble Prediction System (MEPS). GEFS is a single-model, global ensemble system comprised of 21 members run at 00, 06, 12, and 18 UTC. In February 2012, NCEP completed an upgrade to GEFS enabling it to run at a horizontal resolution of T254 (approximately 55km) for hours 0 through 192 and T190 (approximately 73km) for hours 198 through 384. The GEFS vertical resolution is L42 for all 16 days. L42 is the shorthand notation used to refer to the number of vertical levels used by the models. The vertical levels in these models are unevenly spaced throughout the column with layers concentrated in the planetary boundary layer and near the tropopause. GEFS initial condition (IC) perturbations are produced by the method of Ensemble Transform with Rescaling (ETR) (Wei et al. 2008) and the Stochastic Total Tendency Perturbation scheme (Hou et al. 2011) used to account for model uncertainties.

GEPS is a multi-model ensemble comprised of 21 members each from NCEP's

Global Forecast System (GFS) and Canadian Meteorological Center's (CMC) Global Environmental Multiscale (GEM), and 20 members from the Fleet Numerical Meteorological and Oceanography Center's (FNMOC) Navy Operational Global Atmospheric Prediction System (NOGAPS), for a total of 62 ensemble members. GEPS is run twice daily on the 00 UTC/12 UTC cycles. GEPS output has one-degree (approximately 110 km) horizontal resolution, and is available at 6-hour time steps out to 240 hours. AFWA creates GEPS forecast products using an ensemble post-processor (Kuchera 2012).

Currently, the CMC GEM grid point model has approximately 66 km horizontal resolution with 74 vertical levels. IC perturbations are created using the ensemble transform Kalman filter (Bishop et al. 2001) along with stochastic physics and perturbed physical parameterizations. The CMC GEM is run twice daily at 00 and 12 UTC out to 384 hours. The FNMOC runs the NOGAPS spectral model in an ensemble forecast system (NOGAPS GEFS). The current horizontal resolution is T159 (approximately 80 km) with a vertical resolution of L42. NOGAPS uses a nine-banded ensemble transform method (McLay and Bishop 2010) to produce its initial condition perturbations and is run twice daily at 00 and 12 UTC out to 384 hours.

AFWA's baseline MEPS is a single-model ensemble system consisting of a 20 km hemispheric domain and a tropical "stripe" around most of the equator. MEPS uses global deterministic models for initial and lateral boundary conditions and is comprised of 10 members created using the Weather Research Forecasting (WRF) framework. MEPS is run twice daily at 06 and 18 UTC out to 144 hours (Kuchera 2012).

It is important to note that model upgrades occur regularly. For the most up-to-date specifications for GEPS and MEPS reference: `https://weather.af.mil/`

confluence/display/AFPUBLIC/Operational+AFWA+Ensemble+Information. For the most up to date information regarding GEFS, CMC GEM, and NOGAPS model specifications refer to http://www.emc.ncep.noaa.gov/GEFS/clog.php, http://weather.gc.ca/ensemble/verifs/model_e.html, and, https://www.fnmoc.navy.mil. The most comprehensive and up-to-date summary of EPS characteristics is at http://www.meted.ucar.edu/nwp/pcu2/ens_matrix/nefs_p1.htm.

2.1.2.1 EPS with Perturbed ICs.

Any difference between the model-estimated and the true state of the atmosphere at the initialization of the NWP model is analysis error. Allen (2009) states that analysis error may result from errors in the observations due to instrument limitations or the inability to observe at all spatial and temporal scales. Leutbecher and Palmer (2008) separate perturbation techniques into two groups: 1) techniques that aim to obtain a sample of initial states from the PDF and 2) techniques that selectively sample initial uncertainty only in those regions that are dynamically the most important for determining the ensemble dispersion.

A technique from the first group is employed by Canada's MSC in their GEM model. The model's initial conditions are created using an ensemble Kalman filter in which assimilated observations are perturbed by pseudo-random numbers within the known uncertainty of the observations. The added noise from this method represents observational error and can be viewed as a feasible approximation of the extended Kalman filter (Leutbecher and Palmer 2008).

An example of a technique from the second group is the bred-vector (BV), which is used at NCEP. A BV is created by adding a random perturbation to the initial state. Both the perturbed and unperturbed states are evolved using the

NWP model over some forecast period (usually on the order of 6 to 24 hours). The vector difference between the original and perturbed states is found and rescaled to match the typical analyses error. This new perturbation is then added to a new state estimate and the procedure is repeated. Finally, after several growth and rescaling steps the BV is obtained (Leutbecher and Palmer 2008; Toth and Kalnay 1997). FNMOC's Ensemble Transform (ET) method is another example of a technique from the second group. Wei et al. (2008) found that "the properties of ET perturbations show that the method resembles breeding, in that they both dynamically cycle the fastest growing non-linear perturbations."

2.1.2.2 EPS with Model Perturbations.

Model error is defined as any difference between the model forecast and the true atmospheric state resulting from the design of the NWP model, including limits in model resolution, mathematical formulation, lateral and surface boundary conditions, and physics (Allen 2009). Leutbecher and Palmer (2008) describe three general methods for representing model error: 1) multi-model ensemble, 2) the perturbed parameter ensemble, and 3) stochastic-dynamic parameterization.

The success of the multi-model ensemble lies in the fact that different models from different institutes have been developed quasi-independently. Therefore, there will be a variety of convective schemes, orographic drag schemes, and numerical approaches within a multi-model ensemble (Leutbecher and Palmer 2008). Mylne et al. (2002) showed that, in addition to choosing NWP systems with similar overall skill, it is also important that the NWP systems are as independent of each other as possible. This maximizes the chance that the strengths of one model will overcome the weaknesses of another model, thus resulting in a more skillful ensemble. The size of a multi-model EPS is limited by the number of models currently avail-

able. In order to overcome this limitation, the perturbed parameter approach was developed (Murphy et al. 2004). The majority of examples involving perturbed parameter ensembles are associated with climate change research.

Unlike the perturbed parameter ensemble, the stochastic-dynamic parameterization approach does not assume that the correct tendency can be given by a deterministic bulk formula. "Indeed, in the stochastic-dynamic approach it is assumed that the assumption of an ensemble of sub-grid processes at any time-step and for any grid box, is a flawed assumption" (Leutbecher and Palmer 2008). Buizza et al. (1999) showed that stochastic physics led to an increase in the spread of the ensemble, improving its performance. An increase in the spread, up to a certain point, allows the probability distribution function (PDF), which is inferred from the distribution of ensemble members' forecasts, to more accurately represent the future state of the atmosphere. As the ensemble spread approaches zero, the ensemble essentially becomes a deterministic forecast.

2.1.2.3 Quantifying EPS Uncertainty.

There are a number of ways to quantify the uncertainty associated with a forecast. Three commonly used methods are the ensemble mean, ensemble spread, and probability of exceedance. The most relevant to this study are the ensemble mean and spread. The ensemble mean is simply the forecast obtained by averaging all the ensemble members' forecasts. The spread is the standard deviation of the ensemble members' forecasts. It represents how much uncertainty exists among the ensemble members. Typically, forecasters employ the ensemble mean and spread with the assumption that the PDF is a normal distribution, which is not necessarily the case (UCAR 2004). Figure 1 depicts the mean (wind barbs) and spread (colored contours) associated with a 250mb wind forecast. The highest spread val-

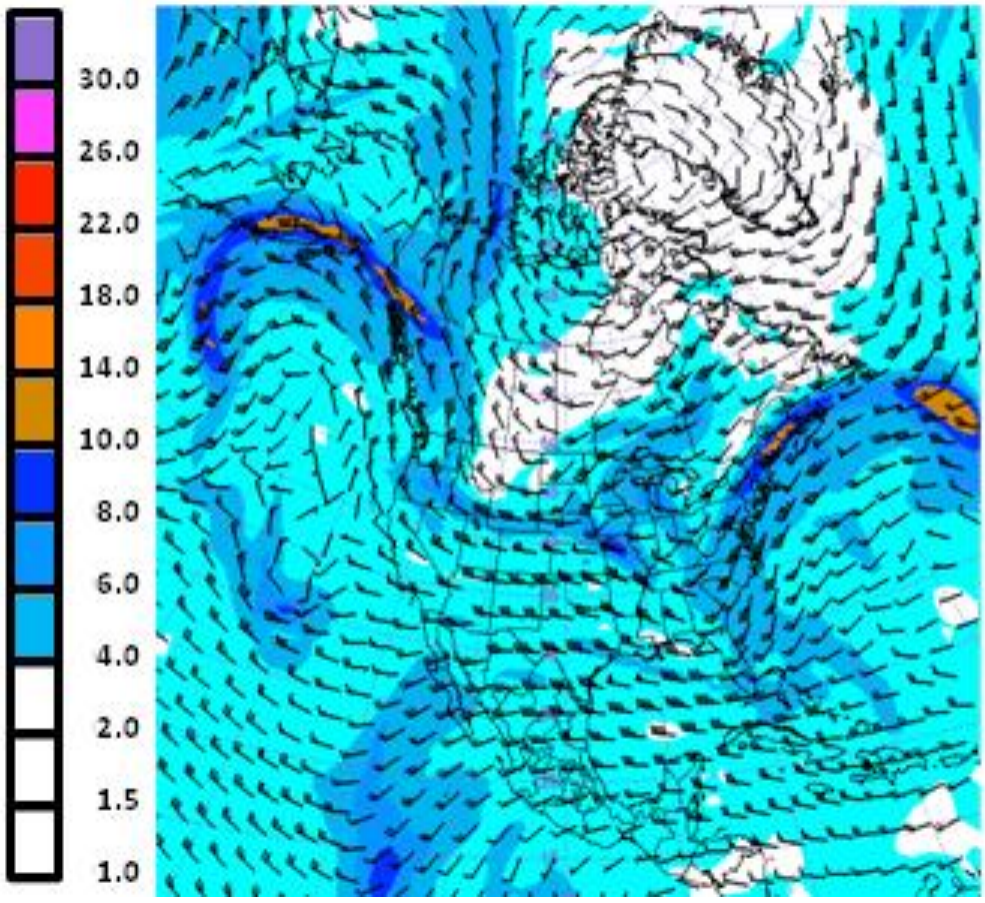

Figure 1. GEFS mean and spread for 250mb winds from the 22 Apr 2013 00 UTC model run, valid 23 Apr 2013 at 00 UTC. Winds are represented by barbs and are in units of ms^{-1}. Each short barb represents 5 ms^{-1}, each long barb 10 ms^{-1}, and pennants 50 ms^{-1}. Spread is represented by the color filled contours with units of ms^{-1}. Image obtained from `http://mag.ncep.noaa.gov`**.**

ues are indicated by the orange shading and are located just south of Alaska, over Northeastern Maine, and east of Maine over the Atlantic Ocean. The white shading located over central Canada and extending northeast over Greenland indicates the lowest spread values.

Figure 2 is an example of a probability of exceedance graphic. The probability of exceedance is based on the fraction of ensemble members that exceed a chosen threshold. The highest probabilities are depicted in red. The locations with the highest probabilities include a large area over the Gulf of Mexico, as well as a small

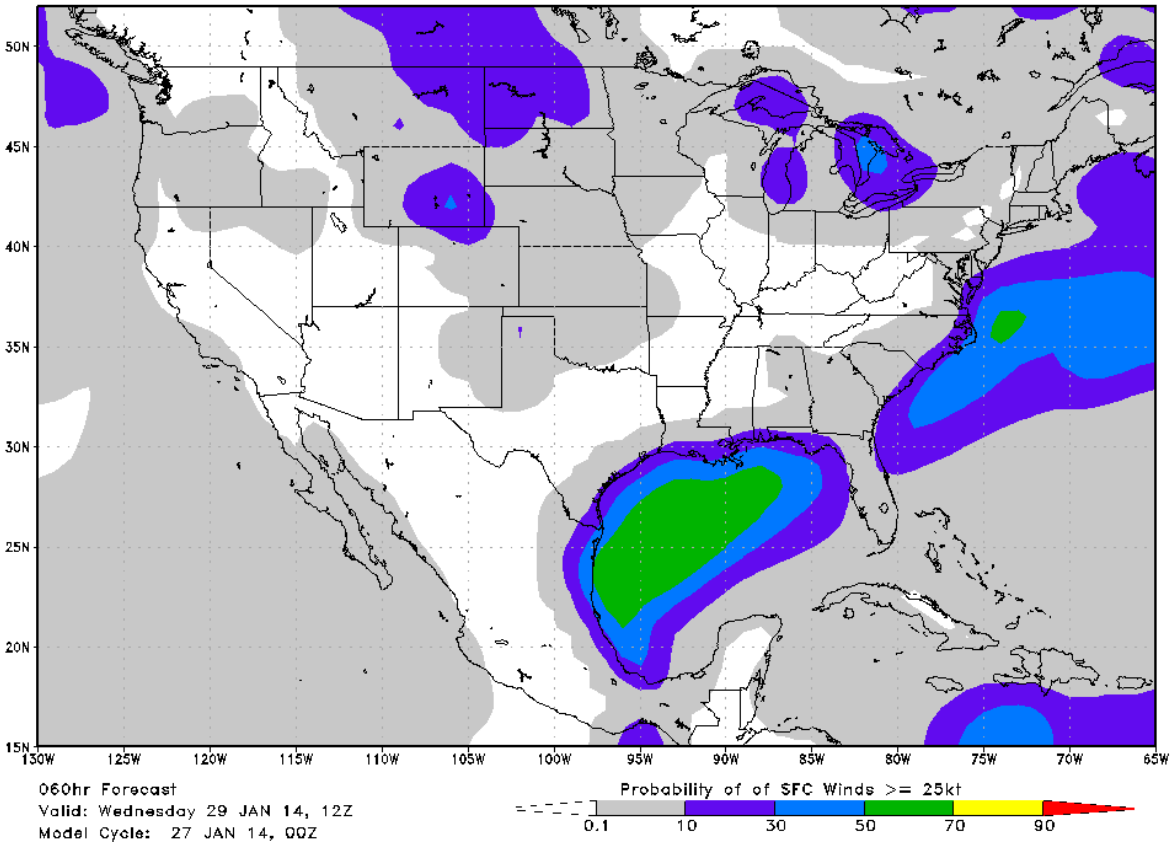

Figure 2. Probability (in percent) of surface winds exceeding 25 kt based on AFWA's Global Ensemble Prediction System (GEPS) valid 12 UTC on 29 Jan 2014. Image obtained from `https://weather.af.mil`.

area in the Atlantic Ocean east of North Carolina. Both of these areas have probabilities ranging from 50-70 percent.

2.2 Advanced Computer Flight Planning System (ACFP)

AMC uses the ACFP system to determine optimal routes of flight based on aircraft performance, winds and temperatures aloft, as well as air traffic control and diplomatic constraints. Additionally, the ACFP system uses detailed performance data for climb, cruise, and descent at all possible aircraft weights. AFWA provides ACFP with GFS model wind, temperature, and geopotential heights for 15 separate atmospheric levels in gridded binary (GRIB) format files every 6 hours.

The GRIB data uses 1° grid spacing and contains output at forecast hours 0-48 at 6-hour intervals and forecast hours 48-96 at 12-hour intervals. In addition to the NWP model data, ACFP also receives climate data from AFWA. Similar to the NWP data, the climate files contain temperature, geopotential heights, and winds for the same 15 atmospheric levels. Climatological wind, temperature, and geopotential height data would be used when the flight plan being generated extends beyond the available model forecast period.

3. Methodology

In this chapter, the methods used to obtain and analyze the data used in this study are explained. The chapter outlines the simplified version of ACFP used to compute forecasted fuel burns. The implementation of the simplifications used for this study allowed it to focus solely on how wind forecast variability affects the forecasted fuel burn for different aircraft and different routes of flight.

3.1 Flight Routes and Aircraft Specifications

In this study, fuel loads planned using GFS deterministic model forecast winds aloft were compared to fuel loads planned using GEFS, GEPS, and MEPS ensemble mean forecast winds for five different pre-determined flight routes and five different aircraft types. The aircraft, routes, and cruise altitudes selected were based on common long-range flights planned by AMC. A summary of the aircraft types, true air speeds (TAS), cruise levels, and fuel burn rates used can be found in Table 1. The average fuel burn error was calculated and analyzed to determine if any model biases were present. The fuel burn root mean square error (RMSE) was also calculated to compare the accuracy of the GFS deterministic model and the three ensembles. Additionally, the forecasted ensemble wind spread was calculated and summed along each flight route to compare the total spread and the fuel burn error.

Five different flight routes representative of routes flown by AMC aircraft were used for the fuel calculations in this study. Table 2 contains the departure and destination latitudes and longitudes and the total great circle distance in nautical miles for each route. It is important to point out that the great circle routes used in this study were a simplification of the more complicated routes of flight actu-

ally utilized by AMC. However, the simplified routes should yield similar outcomes in the fuel burn error to the more complicated routes since the deviations in the more complicated routes account for a small percentage of the total distance, and thus total flight time. The chosen flight routes covered a range of different meteorological phenomena, creating the opportunity for variability between the GFS deterministic and the three ensemble mean forecasts. For example, the flight levels at which the C-17, C-5, KC-10, and KC-135 cruise ensured that the aircraft passed through the main core of the the polar front jet stream at least once and possibly twice, during the Ramstein AB, Germany to Dover AFB, DE flight. The same can be said for the Joint Base Lewis-McChord, WA to Yokota AB, Japan route. The Travis AFB, CA to Manas AB, Kyrgystan fight route encountered the jet stream twice, once as the aircraft traveled north toward the pole and again after it crossed the pole and traveled south over western Asia. The Charleston AFB, SC to Travis AFB, CA route crossed the Rocky Mountains, another source of variability in the upper-level wind forecasts. Finally, the Travis AFB, CA to Joint Base Pearl Harbor-Hickam, HI route crossed from prevailing mid-latitude westerly winds into tropical easterlies, and in the process, encountered the sub-tropical jet stream.

Table 1. Aircraft Specifications

Aircraft	True Air Speed (kt)	Cruise Level (mb)	Cruise Level (ft)	Fuel Burn Rate (lb/hr)
C-130	300	500	FL180	5000
C-17	443	250-400	FL340-240	18000
C-5	450	250-400	FL340-240	25000
KC-10	470	250-400	FL340-240	18000
KC-135	380	250-400	FL340-240	10000

Figure 3 depicts a mid-latitude Pacific great circle flight path originating from Joint Base Lewis-McChord, WA (KTCM; 47°08'16"N 122°28'35"W) with a destination of Yokota AB, Japan (RJTY; 35°44'55"N 139°20'55"E). Figure 4 shows the great circle flight path for a mid-latitude Atlantic flight from Ramstein, AB,

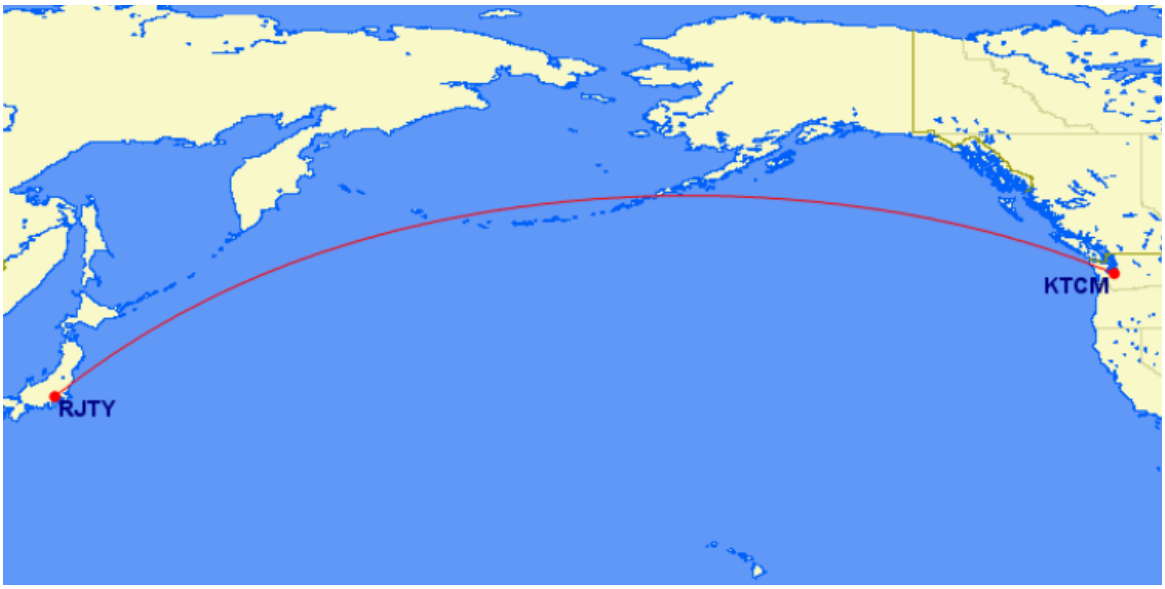

Figure 3. Projected great circle flight path for a mid-latitude Pacific flight departing from McChord AFB, WA and arriving at Yakota AB, Japan. Figure created using `http://www.gcmap.com`**.**

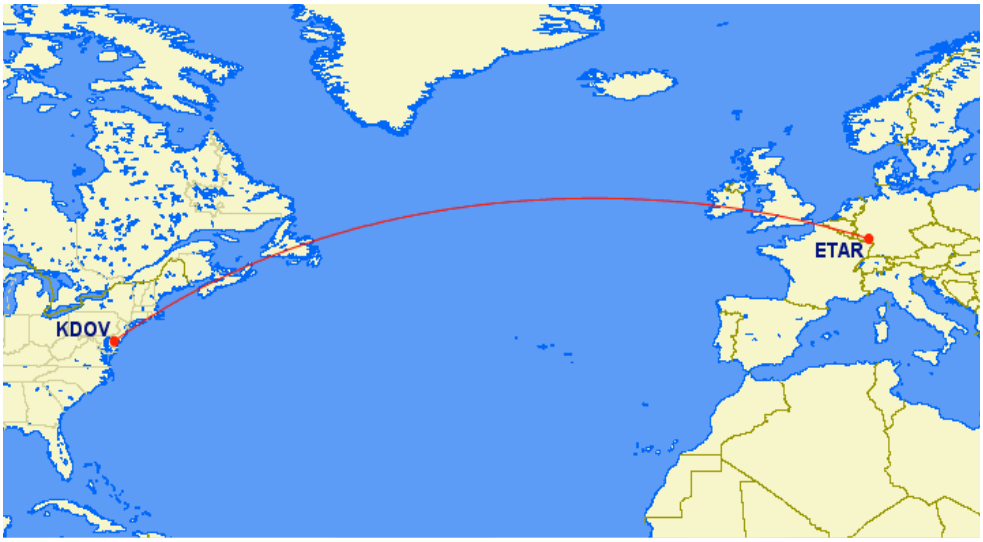

Figure 4. Projected great circle flight path for a mid-latitude Atlantic flight departing from Ramstein AB, Germany and arriving at Dover AFB,DE. Figure created using `http://www.gcmap.com`**.**

Table 2. Flight Routes

Flight Route	Depart lat/lon (degrees)	Dest lat/lon (degrees)	Great Circle Dist (nm)
KTCM-RJTY	47.138N/122.476W	35.749N/139.349E	4179
ETAR-KDOV	49.437N/7.600E	39.129N/75.466W	3444
KSUU-PHIK	38.265N/121.924W	21.319N/157.922W	2117
KCHS-KSUU	32.899N/80.041W	38.265N/121.924W	2055
KSUU-UCFM	38.265N/121.924W	43.061N/74.478E	5850

Germany (ETAR; 49°26'13"N 7°36'01"E) to Dover AFB, DE (KDOV; 39°07'46"N 75°27'57"W). Figure 5 shows the flight path for a low-latitude Pacific route from

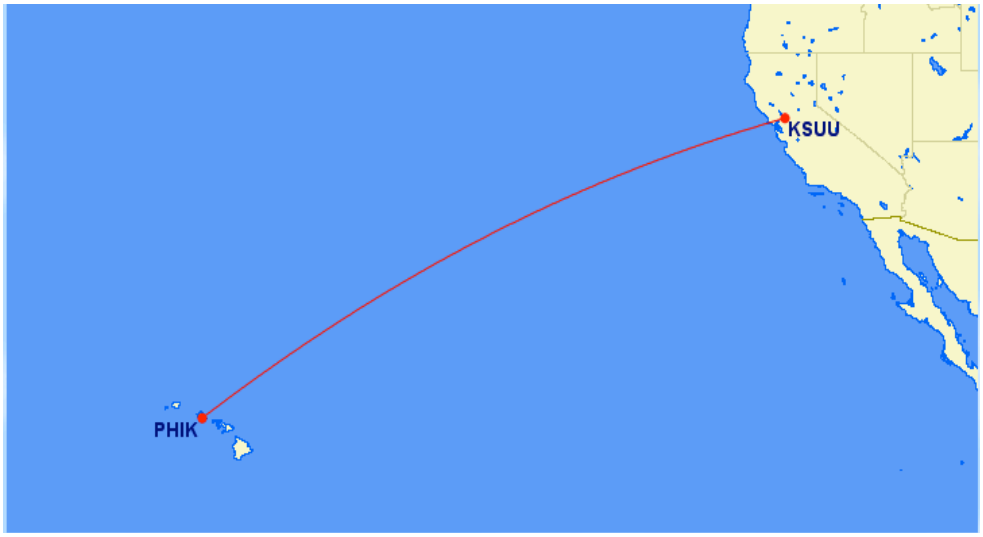

Figure 5. Projected great circle flight path for a low-latitude Pacific flight departing from Travis AFB, CA and arriving at Hickam AFB, HI. Figure created using `http://www.gcmap.com`**.**

Travis AFB, CA (KSUU; 38°15'52"N 121°55'27"W) to Joint Base Pearl Harbor-Hickam, HI (PHIK; 21°19'07"N 157°55'2"W). Figure 6 depicts a lower mid-latitude CONUS flight path from Charleston AFB, SC (KCHS; 32°53'55"N 80°02'26"W) to Travis AFB, CA (KSUU; 38°15'52"N 121°55'27"W). Figure 7 shows an over the pole flight originating from Travis AFB, CA (KSUU; 38°15'52"N 121°55'27"W) and arriving at Manas AB, Kyrgyzstan (UCFM; 43°03'41"N 74°28'39"E). For the sake of consistency in relating positive/negative fuel burn errors to under-/over-forcasted westerly winds, the fuel burn was calculated for flights traveling from east to west

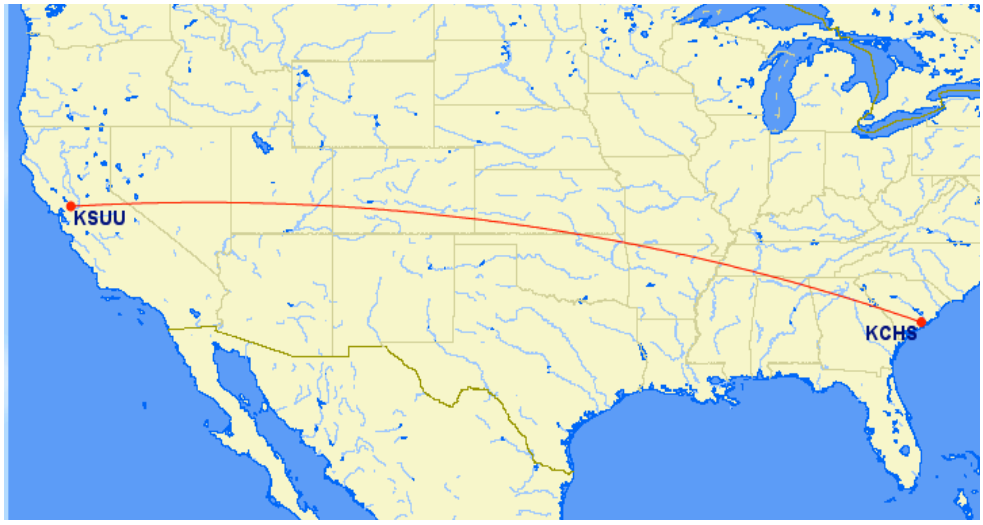

Figure 6. Projected great circle flight path for a lower mid-latitude CONUS flight departing from Charleston AFB, SC and arriving at Travis AFB, CA . Figure created using `http://www.gcmap.com`.

along each route. Therefore, in this study, negative fuel burn errors are always associated with the headwind being under-forecast while positive fuel burn errors are associated with the headwind being over-forecast.

3.2 Required data

The GFS deterministic model and the GEFS, GEPS, and MEPS ensembles were used in this study. The u (east-west) and v (north-south) wind components at pressure levels 500mb, 400mb, 300mb, and 250mb were obtained from the GFS model. The u and v components of the ensemble mean wind and spread at 500mb and 250mb were obtained from the GEFS and GEPS while the u and v components of the ensemble mean wind and spread at pressure levels 500mb, 400mb, 300mb, and 250mb were obtained from the MEPS. Note that the 300mb and 400mb levels were not used for the GEFS and GEPS because those levels are not part of the standard output from those ensembles. Both the GFS and GEFS are run at 6-hour intervals at 00, 06, 12, and 18 UTC daily. GEPS and MEPS are run at 12-hour in-

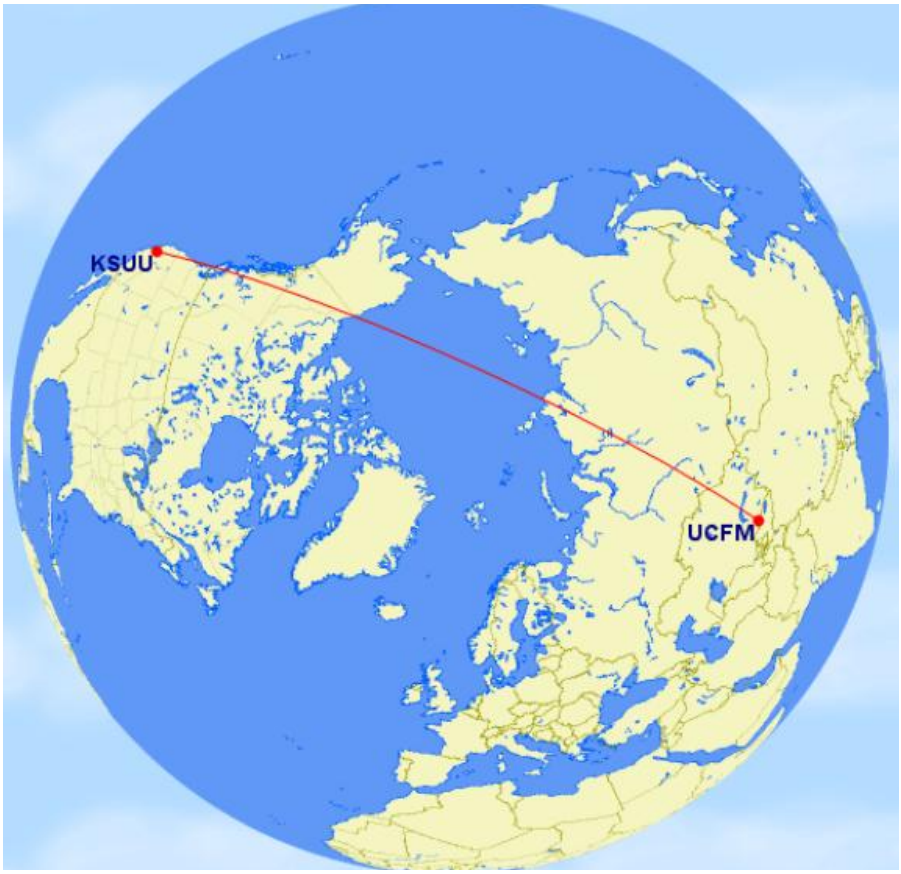

Figure 7. Projected great circle flight path for an over the pole flight departing from Travis AFB, CA and arriving at Manas AB, Kyrgyzstan. Figure created using `http://www.gcmap.com`**.**

tervals. GEPS runs occur at 00 and 12 UTC while the MEPS runs occur at 06 and 18 UTC. Both the GEPS and MEPS data sets were obtained from AFWA. For each model run the forecast fuel burn was computed for each route, each aircraft type, and each cruise flight level, for each forecast hour from 0 through 120 at 6-hour intervals.

3.3 Calculating fuel burn

In addition to wind, air temperature, aircraft weight, and altitude all affect the amount of fuel an aircraft burns throughout the flight. In order to eliminate these other sources of variability in fuel burn calculations, this study: 1) employed a pre-

set fuel burn rate for each type of aircraft based on a representative aircraft weight and altitude for each route; 2) neglected fuel burn during take-off, climb, descent and landing; 3) neglected temperature variation; 4) employed a pre-determined TAS for each aircraft type; and 5) assumed a constant aircraft weight throughout the flight. The fuel burn rate, TAS, and cruise altitudes for each of the five aircraft were provided by AMC.

An iterative algorithm was used to calculate the total fuel burn along each of the great circle routes specified in Table 2. As a hypothetical aircraft progressed along the route its ground speed, heading and position was re-calculated each minute using the wind data from the nearest model grid point. The nearest grid point was used, rather than a weighted average of nearest neighbors, in order to produce manageable run times. This re-calculation took place every 1 minute. The 1-minute time step was chosen in order to capitalize on the full resolution of the model by ensuring that the aircraft did not travel past the next grid point before an updated calculation of the ground speed, heading, and position took place. A 1-minute time step used with the TAS of the fastest aircraft, the KC-10 (470 kt or 242 ms^{-1}), equates to the aircraft traveling approximately 8 nm min^{-1} or 12.5 km min^{-1}. This distance is less than the 20 km grid spacing of the MEPS, the highest resolution ensemble. The 1-minute updates continue until the aircraft reaches its destination. The total time of flight was then used to calculate the forecasted fuel burn using the constant fuel burn rate for each aircraft.

3.4 Model comparison

Once the forecasted fuel burn was calculated using the deterministic forecast and ensemble mean at each forecast hour, the fuel burn error was calculated using the +00 hour forecast (i.e. the analysis) for a particular date/time as the ground

truth for verification of all previous forecasts valid at that date/time. For example, the 19 September 18 UTC +00 hr forecast was used to verify the 19 September 12 UTC +06 hour forecast, the 19 September 06 UTC +12 hour forecast, and the 19 September 00 UTC +18 hour forecast. This method of verification was chosen because it is the common practice among NWP researchers (Joint Working Group on Forecast Verification Research 2012). The fuel burn error was calculated by subtracting the actual fuel burn (ground truth) from the forecasted fuel burn. Therefore, a negative fuel burn error indicates not enough fuel planned for the flight, while a positive fuel burn error indicates excess fuel planned. Once all flights were analyzed, the average fuel burn error and RMSE for each forecast hour was calculated using Equations (1) and (2) respectively:

$$\overline{FB}_{error} = \frac{1}{N}\sum_{i=1}^{N}\left(FB_{Fcst} - FB_{Actual}\right)_i \tag{1}$$

$$FB_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(FB_{Fcst} - FB_{Actual})_i^2} \tag{2}$$

where N is the sample size, FB_{Actual} is the ground truth fuel burn, and FB_{Fcst} is the forecasted fuel burn.

3.5 Ensemble Spread vs. Fuel Burn Error Correlation

In order to answer the second question, whether or not a correlation exists between the ensemble forecast uncertainty and error in the planned fuel burn, the total ensemble wind forecast spread was calculated using a similar process to that used to calculate the fuel burn. Instead of using the mean wind value at the nearest grid point, the ensemble spread was used. The u and v components of the ensemble spread were used to calculate a "combined wind spread" at each model grid point

using Equation (3).

$$spread_c = \sqrt{spread_u^2 + spread_v^2} \tag{3}$$

The combined wind spread was then summed over the total distance of the flight route to yield an additive wind spread. The additive wind spread for forecast hours 06 through 120 for GEFS and 12 through 120 for GEPS and MEPS was then compared to the absolute value of the fuel burn error for the same forecast hours to determine correlation using Equation (4):

$$r_{xy} = \frac{1}{1-N} \sum \left(\frac{x - \bar{x}}{s_x} \right) \left(\frac{y - \bar{y}}{s_y} \right) \tag{4}$$

In Equation (4), x is the fuel burn error, $\bar{x}$ is the average fuel burn error, y is the additive wind spread, $\bar{y}$ is the average additive wind spread, s_x is the standard deviation of the fuel burn error, and s_y is the standard deviation of the additive wind spread. N remains the sample size.

4. Results and Analysis

This chapter presents the results obtained during this study. The first section compares the deterministic and ensemble mean wind forecasts through analysis of the average fuel burn error and RMSE for each aircraft and route of flight. The second section presents results to determine correlation between the ensemble spread and the fuel burn error.

This study looked at 65 different flight route and aircraft combinations for one deterministic model and three ensembles. The date range for all four prediction systems is 19 September through 17 November 2013. Average fuel burn error results are presented for only FL180 (500mb) and FL340 (250mb) because those are the only levels at which all three EPSs produce output and because the results at both FL240 (400mb) and FL300 (300mb) were similar. All four flight levels are discussed in a sub-section on the RMSE results. The EPSs typically out-performed the GFS deterministic forecast for the early forecast hours (f12 to f48) at FL180 and FL340, however, results were not as favorable at FL240 and FL300 when the only ensemble available for comparison was the MEPS.

4.1 GFS Deterministic vs. Ensemble Mean Forecasts

4.1.1 Average Fuel Burn Error.

The first goal of this study was to determine if ensemble mean wind forecasts are superior to deterministic wind forecasts when planning fuel burn loads for long-range airlift flights. The first result analyzed was the average fuel burn error, which shows forecast performance trends from hours 06 to 120 and provides insight into whether the deterministic model and/or ensembles have a bias in upper level wind forecasts.

With the exception of the ETAR-KDOV flight route, each different flight route and aircraft combination shows a similar trend in which the average fuel burn error becomes increasingly negative as the forecast hour increases. This trend indicates a slight bias in both the GFS deterministic model and all three ensembles. Given that all five routes of flight travel east to west, this bias indicates that each of the four prediction systems has a tendency to under-forecast the westerly winds aloft (i.e. headwinds).

Two examples are presented here. The first shows the highest range of variability for the GFS deterministic model and the GEFS, GEPS, and MEPS ensembles. The large range between positive and negative across all forecast times was isolated to the ETAR-KDOV route. The second case shows the average fuel burn error being negative at all forecast hours for the GFS deterministic and the GEFS, GEPS, and MEPS ensembles. This case more closely represents the remaining aircraft/flight route combinations used in this study as they all showed the majority of the forecast hours having negative average fuel burn errors.

Figure 8 shows the average fuel burn error for a C-130 departing from Ramstein AB, Germany (ETAR) and landing at Dover AFB, DE (KDOV). The GFS deterministic model is represented by a black square, GEFS a red cross, GEPS a blue asterisk, and MEPS a magenta diamond. The GFS deterministic model, GEFS, and GEPS ensembles all begin with an average fuel burn error near zero at forecast hour 12 (f12). It is important to note that an average error of zero does not mean the model was perfect; it merely means that roughly equal numbers of positive and negative fuel burn errors averaged out to near zero. Both the GFS deterministic and GEFS average fuel burn errors remain near zero until f42, when the error increases to 80 and 58 pounds respectively. From this point the GFS deterministic average fuel burn error increases at a faster rate than the GEFS. The error plateaus

near 200 pounds at f60 and remains there until f90. From f96 until the end of the period the GFS average fuel burn error increases to a maximum of 385 pounds. The GEFS average fuel burn error has a maximum of 100 pounds at f60 but decreases to near zero by f84. The final 30 hours of the forecast period the GEFS average fuel burn errors are negative. The GEPS average fuel burn errors are negative for all forecast hours and range from -15 pounds to -96 pounds. The MEPS average fuel burn errors also remain negative for all forecast hours, however the range of errors is much larger with a minimum error of -27 pounds at f24 and a maximum error of -490 pounds at f120.

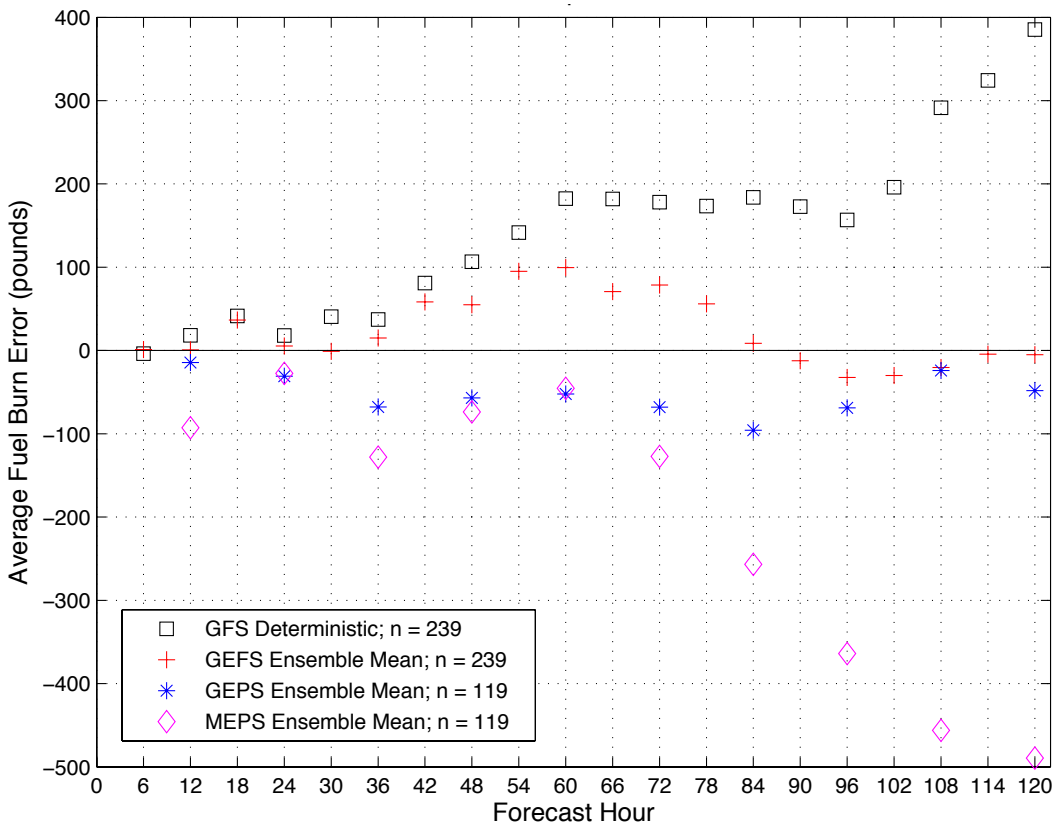

Figure 8. Average fuel burn error in pounds as a function of model forecast hour for a C-130 at cruise level 500mb (FL180) and flight route ETAR-KDOV. The timeframe of the data is 19 Sep – 17 Nov 2013. This case shows the most variability across all forecast hours. Average fuel burn error is the actual fuel burn subtracted from the forecasted fuel burn, thus negative values mean the planned fuel was under-estimated based on the forecasted winds along the flight route and positive values indicate the planned fuel was over-estimated.

This particular route and aircraft combination is unique in that is shows the GFS model with all positive average fuel burn errors, which indicates the model is over-forecasting the head wind, while the MEPS and GEPS both have negative average fuel burn errors for all forecast times. The GEFS shows a combination of both positive and negative errors with a slight bias toward over-forecasting the headwind.

These biases were further investigated to estimate their magnitude. This was done by using the true air speed and fuel burn rates from Table 1 and the flight route distances from Table 2. The actual wind bias values for the GFS, GEFS, GEPS, and MEPS at f36 are 0.14 ms^{-1} , 0.05 ms^{-1} , 0.14 ms^{-1} , and 0.5 ms^{-1} respectively while the bias values for the GFS and MEPS at f120 increase to 1.1 ms^{-1} and 1.4 ms^{-1} respectively. The GEFS and GEPS bias values remain very similar to their values at f36 throughout the range of forecast hours. These wind bias magnitudes are small in comparison with the uncertainty in wind observations themselves, which is between 1 ms^{-1} and 2 ms^{-1}.

The C-130 ETAR-KDOV exhibits a much larger range than the KSUU-PHIK and KCHS-KSUU routes (results not shown) between the GFS and the GEFS, GEPS, and MEPS for forecast hours 48 and beyond (particularly the GFS and MEPS). This may be attributed to individual performance differences between the GFS deterministic model and the MEPS at mid-latitudes versus tropics.

The average fuel burn error results for the GFS deterministic model, GEFS, GEPS, and MEPS for a KC-135 departing from Joint Base Lewis-McChord, WA (KTCM) and landing at Yokota AB, Japan (RJTY) are shown in Figure 9. The overall trend for this flight route and aircraft type shows the average fuel burn error becoming more negative over time. The KCHS-KSUU, KSUU-PHIK, and KSUU-UCFM routes show a very similar trend. The only exception is that for these three

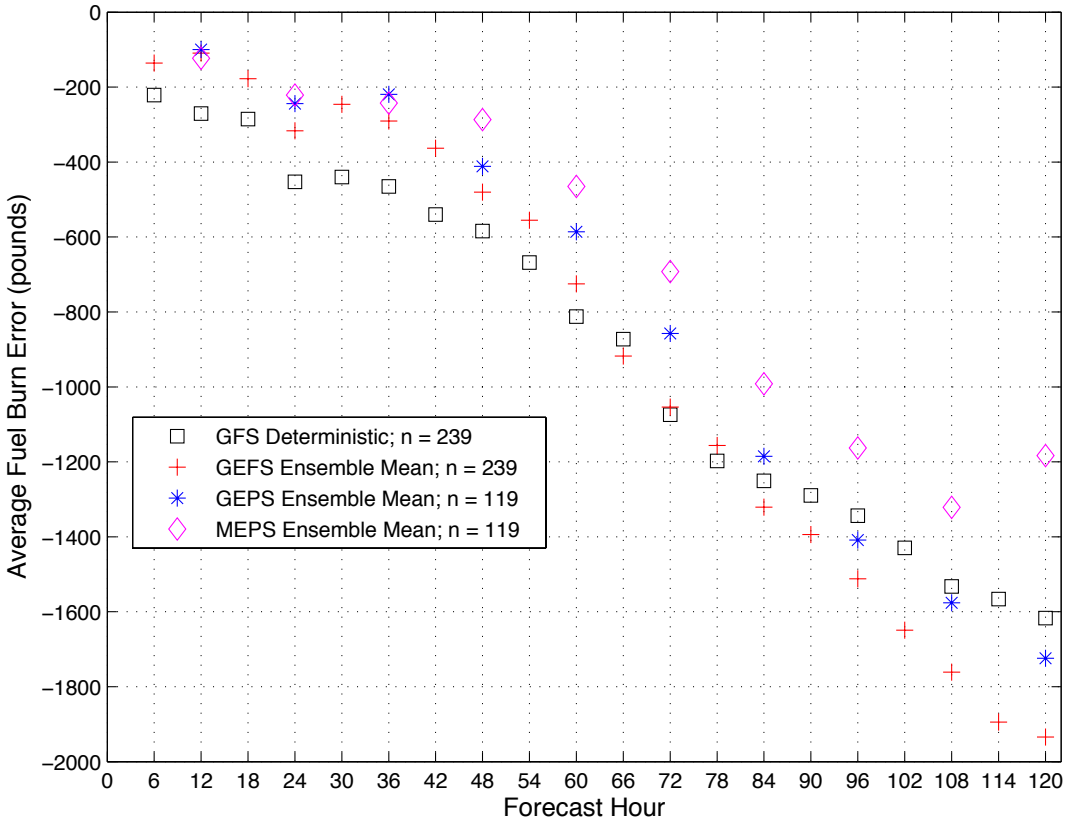

Figure 9. Average fuel burn error in pounds as a function of model forecast hour for a KC-135 at cruise level 250mb (FL340) and flight route KTCM-RJTY. The time-frame of the data is 19 Sep – 17 Nov 2013. This case shows a negative bias which indicates the deterministic and all three ensembles are under-forecasting the head-wind.

routes there is typically at least one prediction system during then early forecast hours (f12 to f36) that had a small positive average fuel burn error.

All three ensembles and the deterministic model again appear to show a bias for this flight route and aircraft type. The range of average fuel burn error doubled from what was shown for the ETAR-KDOV route (Figure 8). This is expected as the fuel burn rate of the KC-135 is twice that of the C-130. When comparing the KC-135 KTCM-RJTY case to all the other flight routes and levels (FL240 and FL300) for the same aircraft, the range of average fuel burn error is consistently the second highest. This is likely due to the lack of surface observations available along the route, which impacts how much variability exists in the deterministic model

and ensemble mean upper-level wind forecasts. These fuel error bias values translate to wind bias values at f36 of 0.98 ms^{-1} , 0.57 ms^{-1} , 0.39 ms^{-1} and 0.39 ms^{-1} for the GFS, GEFS, GEPS, and MEPS respectively, while bias at f120 increased to 2 ms^{-1} , 3.9 ms^{-1} , 2.9 ms^{-1} , and 1.9 ms^{-1} respectively. These wind forecast bias values at f36 were again small relative to wind observing accuracy, while at f120 the bias values for GEFS and GEPS were slightly larger than the accuracy of wind observations.

4.1.2 Fuel Burn Root Mean Square Error (RMSE).

A more significant fuel burn error metric is the RMSE, which quantifies the accuracy of a model. While 65 different aircraft/route combinations were tested, a total of 17 figures are presented here to represent the results. One consideration used when selecting the aircraft and flight route combinations was the feasibility of a non-stop flight for the aircraft. Using this criterion, the KTCM-RJTY and KSUU-UCFM routes for the C-130 were eliminated. For the remaining four aircraft, all five routes of flight were analyzed and the best and worst case, in terms of RMSE reduction by the ensembles, were selected to represent all cases tested. In addition to the figures, a table is used to summarize the 36-hr RMSE for the GFS deterministic, GEFS, GEPS, and MEPS for all five routes of flight and all five aircraft at their respective cruise levels. Additional tables with the RMSE values for forecast hours 12, 24, 60, 84, and 120 can be found in Appendix A. In Figures 10 through 26, the forecast hour is depicted on the x-axis, while the RMSE, in pounds of fuel, is depicted on the y-axis. The GFS deterministic model and the GEFS, GEPS, and MEPS are represented using the same shapes and colors that were used in Figures 8 and 9. This analysis will address the first objective of this study.

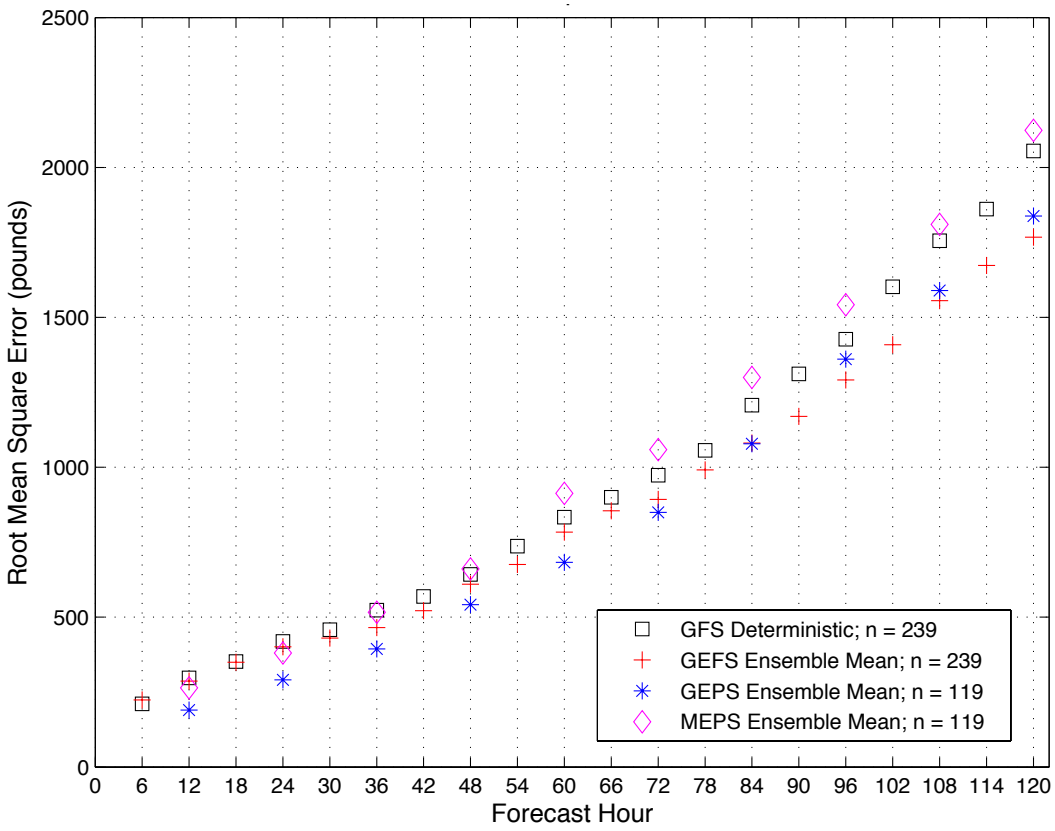

Figure 10. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-130 at cruise level 500mb (FL180) and flight route ETAR-KDOV. The timeframe of the data is 19 Sep – 17 Nov 2013.

4.1.2.1 ETAR-KDOV.

Figure 10 depicts the root mean square error of the fuel burn calculations for a C-130 flying from Ramstein to Dover. The overall trend of the GFS deterministic, GEFS, GEPS, and MEPS, shows that RMSE increases with the forecast hour which is to be expected since forecast accuracy decreases as forecast time increases. The GEFS and GEPS are more accurate than the GFS deterministic for all forecast hours, while the MEPS is only more accurate during forecast hours 12 to 36. The percentage values discussed throughout this section were calculated by subtracting the ensemble RMSE from the GFS deterministic RMSE and dividing the difference by the GFS deterministic RMSE. At f12 the GEPS shows the largest im-

provement over the GFS deterministic model at 36%. The MEPS and GEFS show 11% and 3% improvement respectively. At f24 both the GEPS and MEPS improvements decrease to 31% and 9% respectively, while the GEFS improvement increases to 4%. This trend continues to f36 where the improvement of GEFS has increased to 11%, while the GEPS and MEPS improvement over the GFS deterministic has decreased to 25% and 1% respectively. By f120, two of the three ensembles (GEFS and GEPS) remain more accurate than the GFS deterministic.

The C-130 had the lowest actual RMSE values when compared to the C-17, C-5, KC-10, and KC-135. The first reason the RMSE values are lower is due to the lower fuel burn rate for the C-130. The second reason is likely based on the C-130's cruise altitude and the location of the polar front jet stream. Unlike the C-17, C-5, KC-10, and KC-135, the C-130 generally flies below the core of the jet stream, therefore it is likely that the errors in the GFS deterministic model, as well as the three ensembles, will be lower, which will yield lower overall RMSEs. The RMSE ranges from 190 to 2124 pounds for the C-130 compared to a range spanning 442 to 6654 pounds for the other four aircraft. The various ranges can be attributed to the different fuel burn rates for each aircraft.

The C-5 flight represents the best case scenario, in terms of ensemble RMSE reduction, as can be seen in Figure 11. Results for this case are similar to the C-130 flight shown in Figure 10. The GEFS and GEPS again remain more accurate than the GFS deterministic for all forecast hours, while the MEPS is more accurate for forecast hours 12 to 36. For this case, the GEPS and MEPS had the largest percentage of improvement over the GFS deterministic for forecast hours 12 to 36. Improvements ranged from 22% to 36% The GEFS was also more accurate than the GFS deterministic during the same three forecast hours, however, the percentages were much lower (ranging from 9% to 18%). Similar to the previous case, RMSE

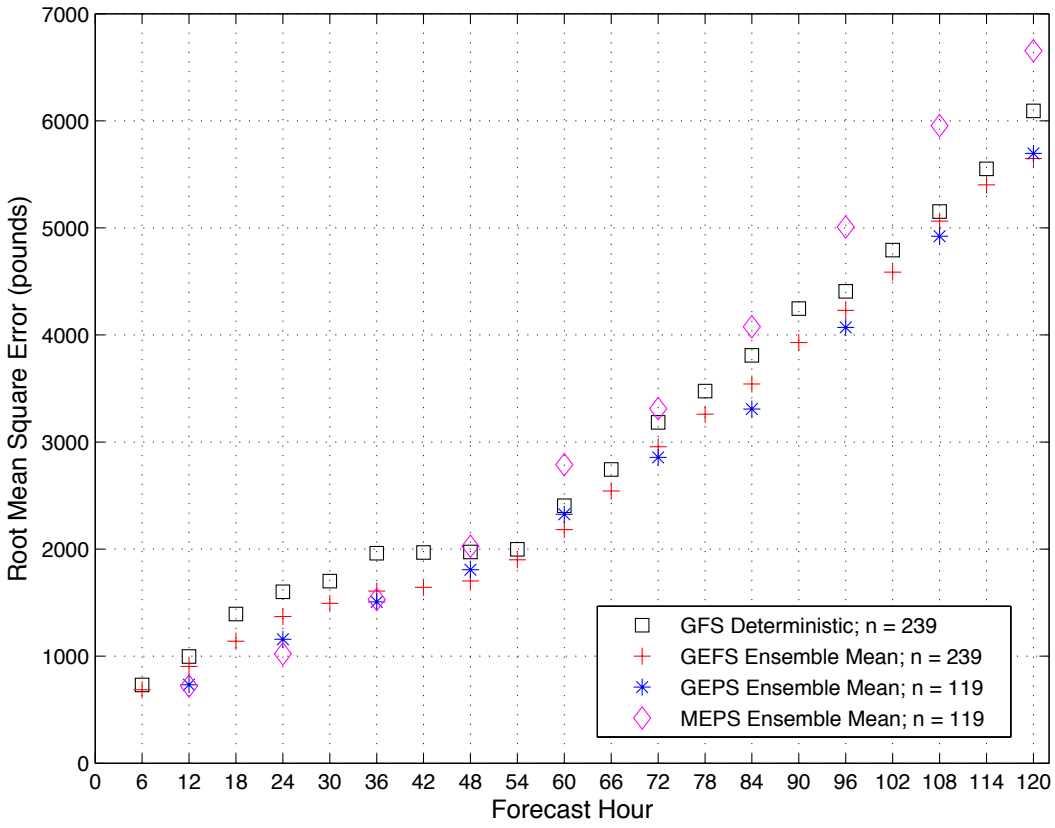

Figure 11. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 250mb (FL340) and flight route ETAR-KDOV. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the best case for this flight route. Best case means the ensembles showed the most improvement over the deterministic model.

increases with time. Both the GEFS and GEPS remain more accurate than the GFS deterministic at f120, but the percentage of improvement has decreased significantly compared to forecast hours 24 and 36, with the GEFS improvement reduced to 9% and the GEPS reduced to 13%. At f120, the MEPS is 9% less accurate than the GFS deterministic model.

The worst case for the ETAR-KDOV route was for the KC-10. Results are shown in Figure 12. Although this case had the smallest difference between the deterministic RMSE and the three ensemble RMSEs, the percentage of improvement shown by the ensembles is very close to the percentages in the best case (Figure 11) because the GFS deterministic RMSE values are the lower. The GEFS improve-

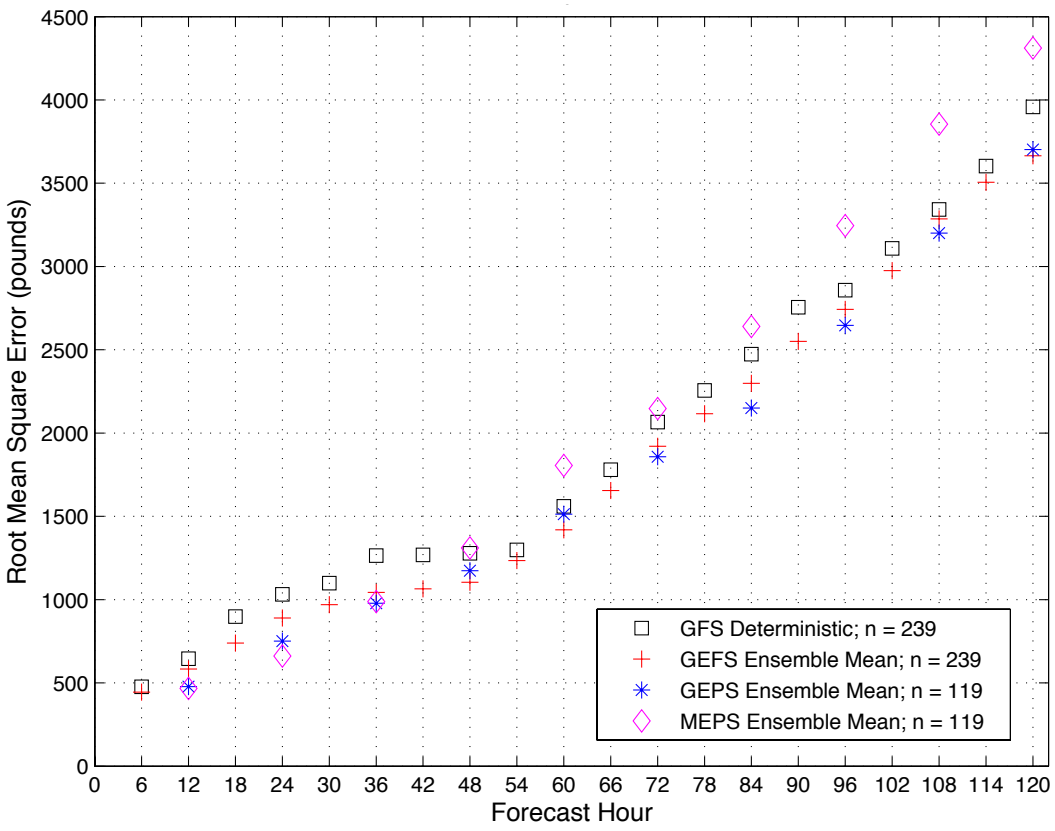

Figure 12. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-10 at cruise level 250mb (FL340) and flight route ETAR-KDOV. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the worst case for this flight route. Worst case means the ensembles showed the least improvement over the deterministic model.

ment ranged from 10% to 17%, the GEPS 23% to 27%, and the MEPS 22% to 38% for forecast hours 12 to 36. The percentage of improvement at f120 remains low for all three ensembles. The GEFS and GEPS are 7% and 6% more accurate respectively while the MEPS is both 9% less accurate than the GFS deterministic model.

4.1.2.2 KCHS-KSUU.

For the Charleston to Travis flight scenario three aircraft were used, the C-130, C-5, and KC-135. As with the ETAR-KDOV route (Figures 10-12) the RMSE increases as the forecast hour increases. The main feature that emerges for this flight route is that for at least two consecutive forecast hours the GFS out-performs all

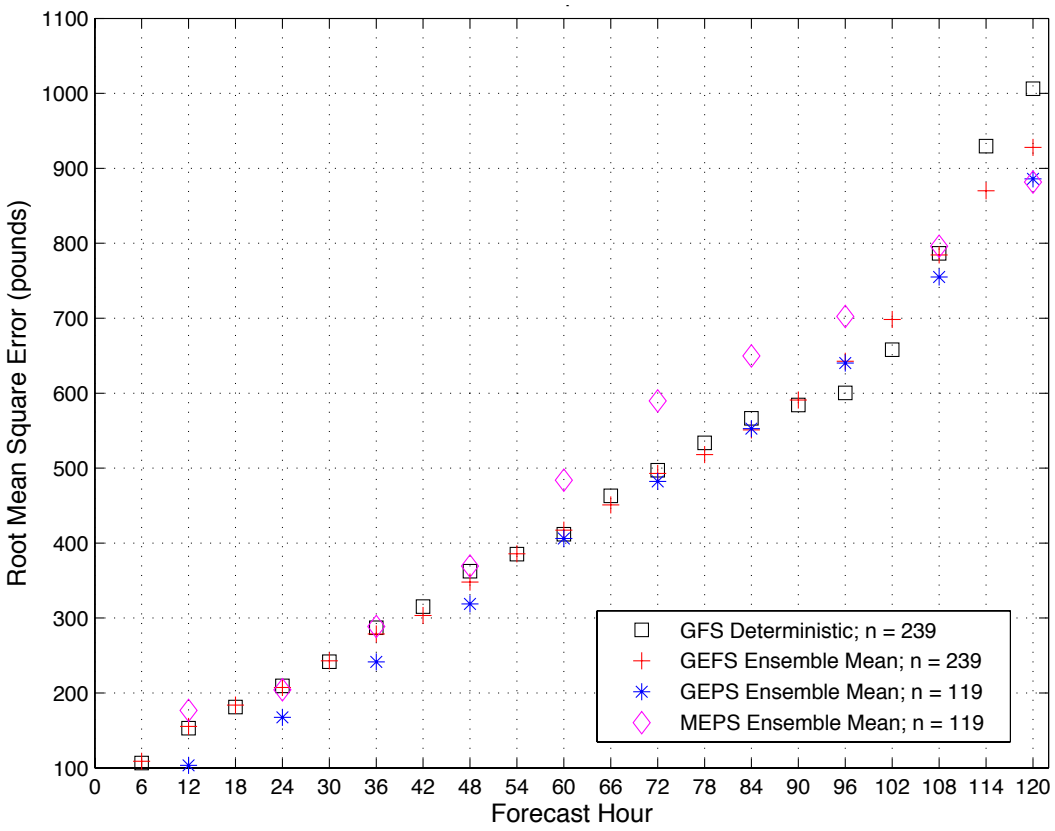

Figure 13. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-130 at cruise level 500mb (FL180) and flight route KCHS-KSUU. The timeframe of the data is 19 Sep – 17 Nov 2013.

three ensembles. These instances all occur after forecast hour 84. Of additional note, the RMSE values for the KCHS-KSUU flight, as a percentage of the improvement over the GFS deterministic model, are lower than the ETAR-KDOV RMSE values. One explanation for this is that the ETAR-KDOV flight path encounters the core of the polar front jet stream, which affords more opportunity for the wind forecasts to be less accurate compared to the lower latitude KCHS-KSUU flight path. Additionally, there are very few surface observations in the Atlantic Ocean, which contributes to a less accurate depiction of the actual state of the atmosphere over that region. A less accurate analysis, might result in even more uncertainty in the upper-level winds throughout time as the error grows.

Figure 13 shows the results for the C-130 flight. At f12 the only the GEPS is

more accurate than the GFS deterministic model, reducing the RMSE by 3%. By f24 all three ensembles are more accurate than the deterministic model, but the percentages of improvement for the GEFS and MEPS remain very low at 1% and 2% respectively. The GEPS improvement over the GFS deterministic has decreased to 20%. Both the GEFS and GEPS remain more accurate than the deterministic model at f36, while the MEPS becomes 1% less accurate. The MEPS remains less accurate than the GFS through f108, but at f120 becomes more accurate than the GFS deterministic model by 12%. Between f36 and f84 the GEFS and GEPS generally remain more accurate than the deterministic model. There is one forecast hour (f96) at which the deterministic model out-performs all three ensembles; by 7% over GEFS and GEPS and 15% over MEPS.

Results for the best case scenario for the KCHS-KSUU route are shown in Figure 14. All three ensembles out-perform the GFS deterministic for forecast hours 12 to 36 and the percentages of improvement range from 2% to 27%. At f48 the GEPS shows only a 1% improvement, the GEFS matches the deterministic model, and the MEPS is less accurate than the GFS by 5%. The GEPS remains 2% to 15% less accurate than the GFS deterministic for forecast hours 60 to 120. During this same time period the MEPS was 7% to 18% less accurate than the deterministic model. For six of the twelve forecast hours between f60 and f120 the deterministic model out-performs all three ensembles.

The worst case for this route, the KC-135, is shown in Figure 15. Similar to the C-5 (Figure 14) results, all three ensembles out-perform the GFS deterministic model during forecast hours 12 to 36. The percentages of improvement for each ensemble are very consistent with the C-5 results as well. At forecast hours 72, 96, and 108 the deterministic model out-performs all three ensembles. The MEPS is the only ensemble that remains more accurate than the deterministic at f120.

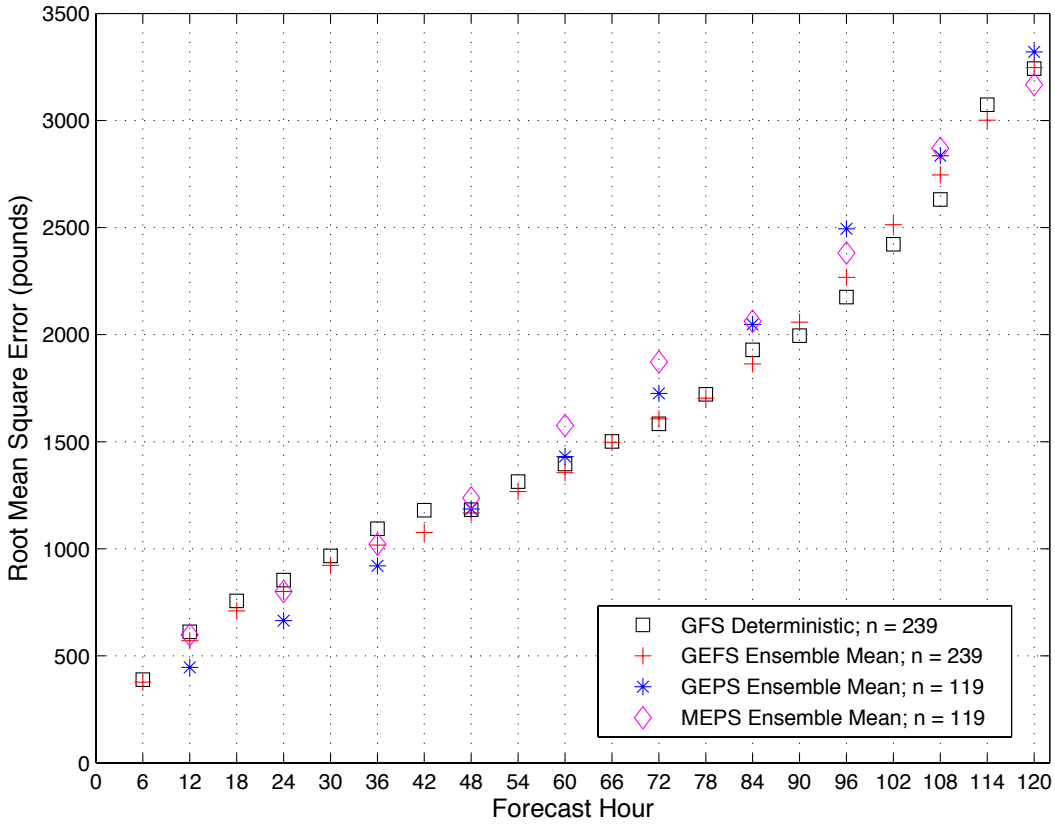

Figure 14. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 250mb (FL340) and flight route KCHS-KSUU. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the best case for this flight route. Best case means the ensembles showed the most improvement over the deterministic model.

4.1.2.3 KSUU-PHIK.

The Travis to Hickam route is the final route for which the C-130 RMSE is analyzed. Results are consistent with those shown for the KCHS-KSUU route (Figures 13–15), with at least one of the three ensembles being out-performed by the GFS deterministic model. Additionally, for all three aircraft, the MEPS performance after f36 is significantly better compared to the ETAR-KDOV and KCHS-KSUU routes. The meteorological features encountered along this route include the transition from mid-latitude westerlies (headwind) to the tropical easterlies (tailwind). Additionally, the jet aircraft will encounter the subtropical jet stream. Since the

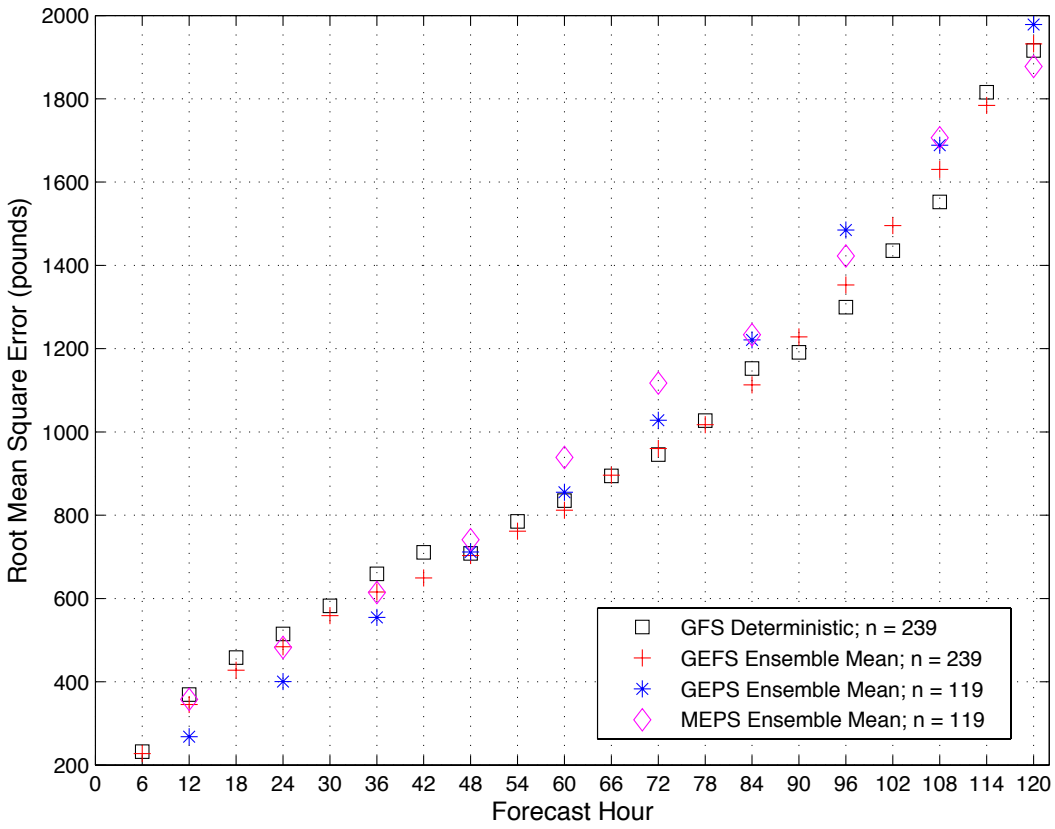

Figure 15. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-135 at cruise level 250mb (FL340) and flight route KCHS-KSUU. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the worst case for this flight route. Worst case means the ensembles showed the least improvement over the deterministic model.

sub-tropical jet is typically weaker than the polar front jet we expect the uncertainty in the wind forecasts to also be lower, producing a lower fuel burn RMSE.

Figure 16 shows that the GEFS and GEPS were more accurate than the GFS deterministic model on only four of the ten 12-hourly forecast times for the C-130. During the most frequently used forecast hours, 12 to 36, the MEPS never out-performed the GFS deterministic model, while the GEPS was more accurate only at f12 (26%) and f24 (11%), and the GEFS was only more accurate (2%) at f24. This route was the first time in which the deterministic model out-performed all three ensembles during one of the first three 12-hourly forecast times (f36). For f48 to f108, the GEPS did not improve or was less accurate than the GFS deterministic

model. The GEFS was slightly better at two of the forecast hours, f72 and f96, and for one forecast hour, f60, it yielded no improvement over the deterministic model. Interestingly, for the second half of the forecast period (f60 to f120) the MEPS remains more accurate than the GFS deterministic model and both the GEFS and GEPS. This is the opposite of what was shown in the ETAR-KDOV (Figure 10) and KCHS-KSUU (Figure 13) results, in which the MEPS was the least accurate during the later forecast hours.

Figure 17 depicts the results for the KSUU-PHIK best case scenario, the C-5. The overall performance of the three ensembles improved during forecast hours 12 to 36. Both the GEPS and MEPS were more accurate than the GFS deterministic model during this time with improvements ranging from 2% to 30%. The GEFS

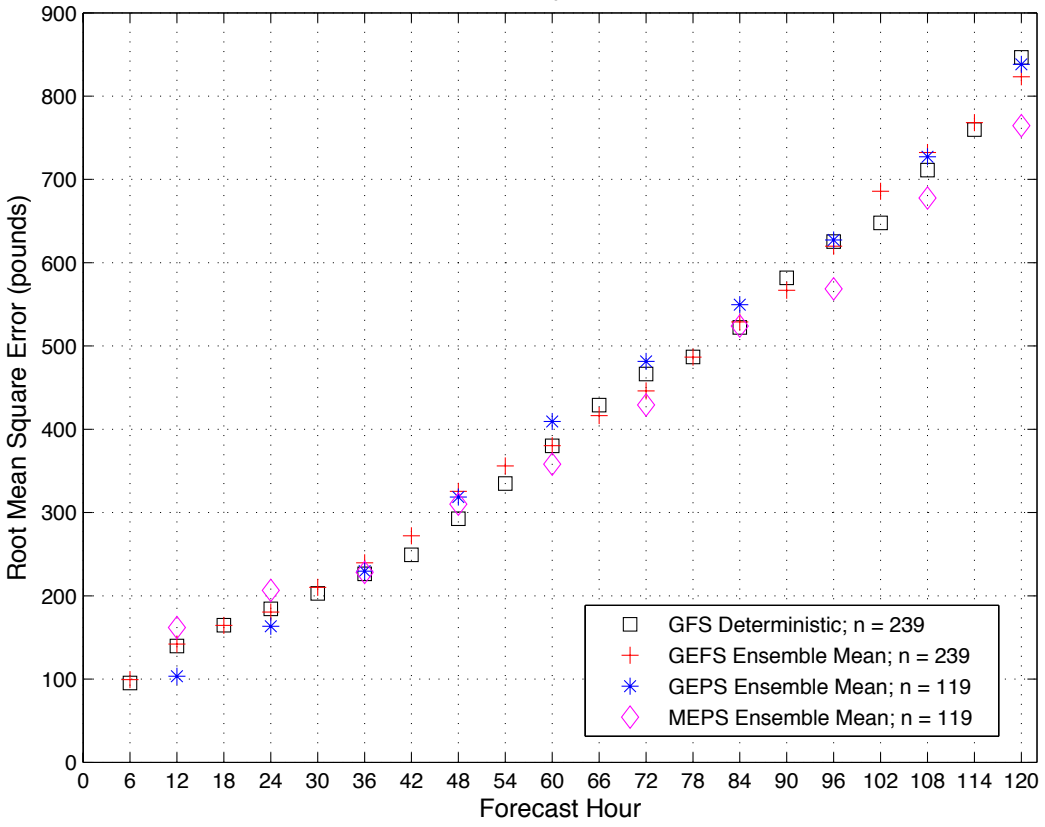

Figure 16. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-130 at cruise level 500mb (FL180) and flight route KSUU-PHIK. The timeframe of the data is 19 Sep – 17 Nov 2013.

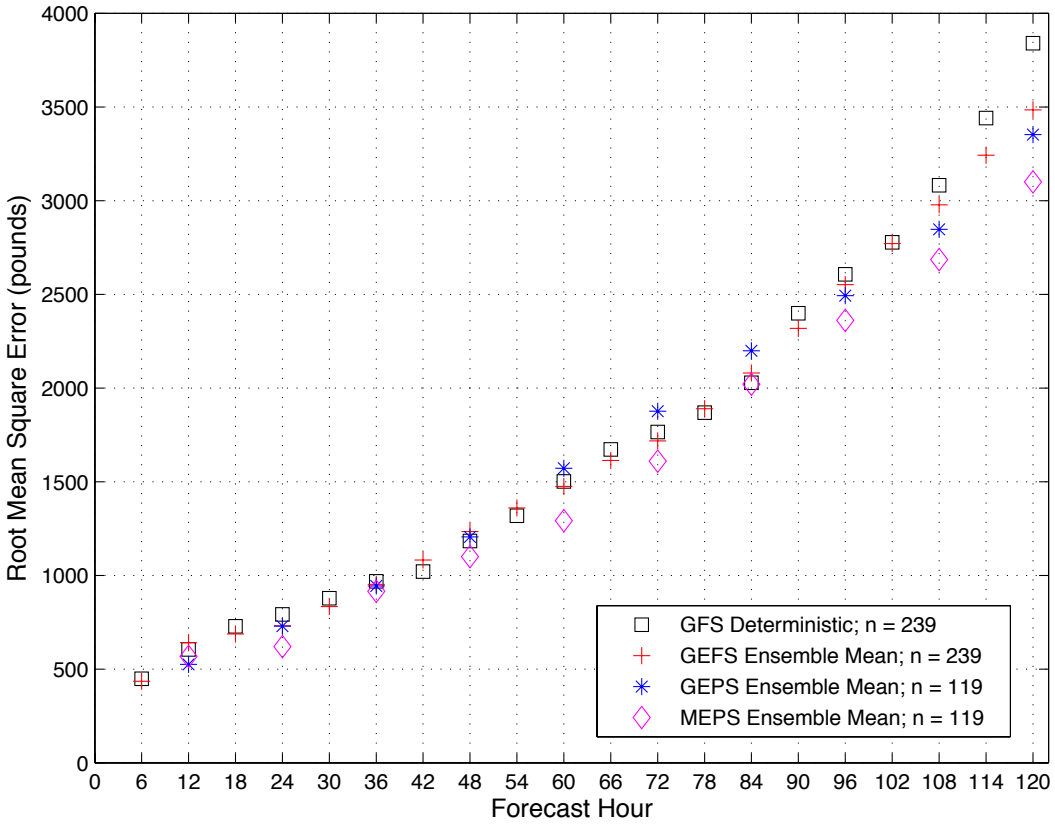

Figure 17. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 250mb (FL340) and flight route KSUU-PHIK. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the best case for this flight route. Best case means the ensembles showed the most improvement over the deterministic model.

was slightly more accurate at f24 and f36 with improvements of 8% and 2% respectively. The GEFS was 6% less accurate than the deterministic model at f12. This was the first case in which the MEPS remained more accurate than the deterministic model for nine of the ten forecast hours and is more accurate than the GEFS and GEPS at all forecast hours except f12.

Results for the KC-135 (worst case) are shown in Figure 18. The overall trends and percentages of improvement for all three ensembles are consistent with the results shown in Figure 17. The MEPS once again was the best performing ensemble, with the lowest RMSE for nine of the ten forecast hours. Additionally, the GEPS out-performed the deterministic model during the early (f12 to f36) and late (f96 to

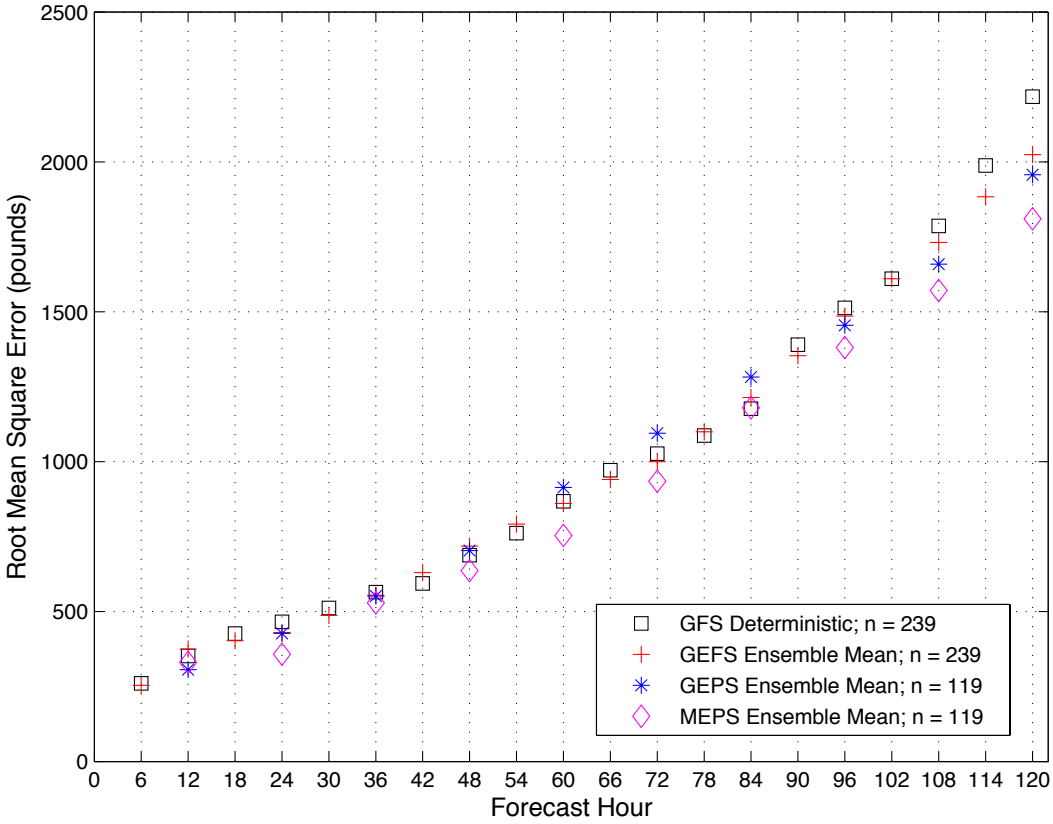

Figure 18. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-135 at cruise level 250mb (FL340) and flight route KSUU-PHIK. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the worst case for this flight route. Worst case means the ensembles showed the least improvement over the deterministic model.

f120) forecast hours, but becomes less accurate during f48 to f84.

4.1.2.4 KSUU-UCFM.

This section discusses the best and worst case results for the Travis to Manas over-the-pole flight. This route is the longest of the five routes tested (5850 nm) and shows the least amount of improvement by the ensembles over the GFS deterministic model. The GFS was more accurate than all three ensembles between forecast hours 36 to 72 for both cases.

The best case results are shown in Figure 19. At f12 both the GEFS and GEPS were more accurate than the GFS deterministic, with the GEFS improvement 3%

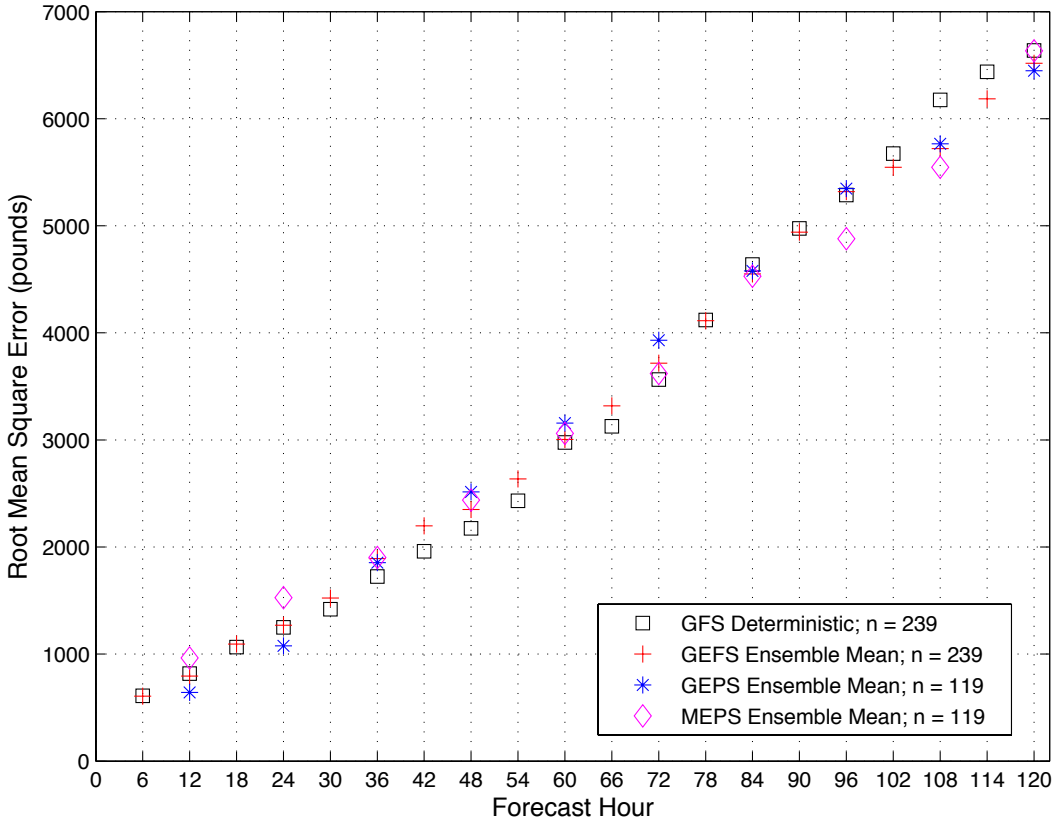

Figure 19. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 250mb (FL340) and flight route KSUU-UCFM. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the best case for this flight route. Best case means the ensembles showed the most improvement over the deterministic model.

and the GEPS 22%. The GEPS was more accurate than the GFS deterministic model at f12 but by f36 the deterministic model has become more accurate than all three ensembles. This trend continued until f84, at which point all three ensembles showed a slight improvement over the deterministic model. For the remainder of the forecast hours at least one of the ensembles was more accurate than the GFS deterministic model, with improvements ranging from 2% to 10%. Overall, the GEPS out-performed the GFS during 50% of the forecast hours, while GEFS outperformed the GFS during 40% of the forecast hours. The MEPS proved to be the least accurate of all three ensembles.

The performance of the GFS deterministic model and all three ensembles for

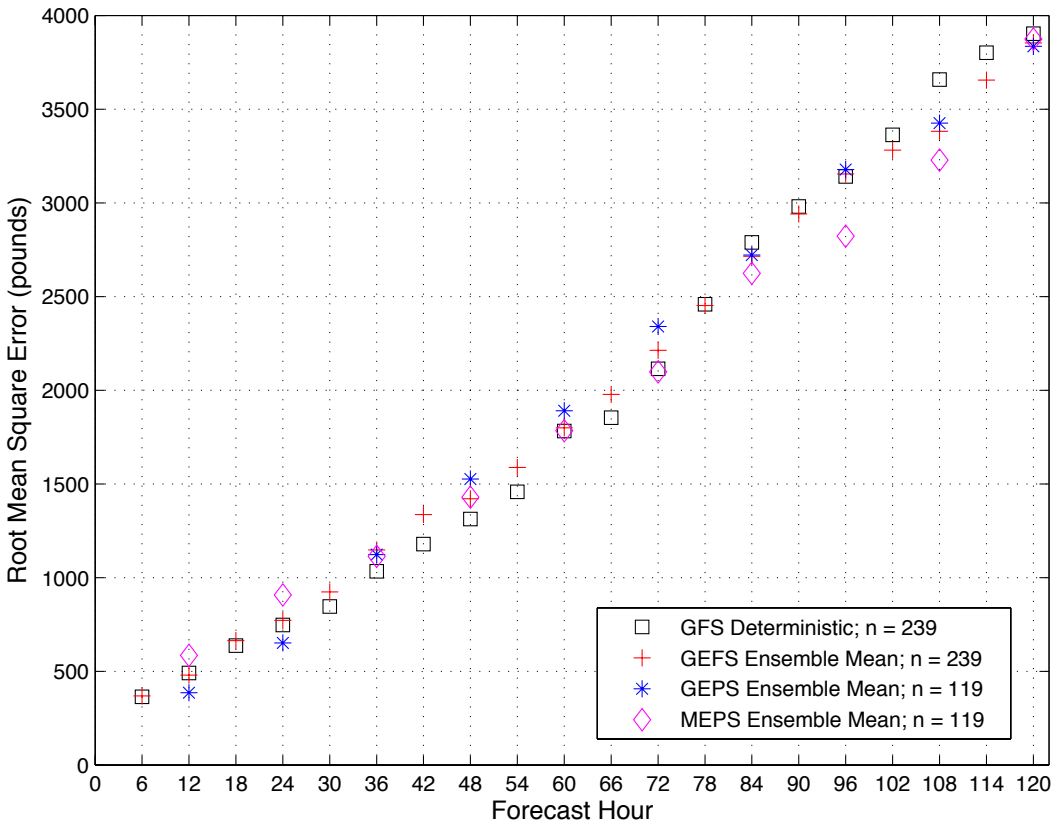

Figure 20. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-135 at cruise level 250mb (FL340) and flight route KSUU-UCFM. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the worst case for this flight route. Worst case means the ensembles showed the least improvement over the deterministic model.

the worst case scenario, the KC-135, remained similar to that of the best case scenario. Results are depicted in Figure 20. For forecast hours 12 to 36 the GEPS was more accurate than the GFS deterministic at f12 and f24. The MEPS remained less accurate than the deterministic model over the same time period. The GFS was the most accurate model during the middle third of the forecast period, but by f72 the MEPS had become more accurate. This occurred 12 hours sooner than during the KSUU-UCFM best case (Figure 19). During forecast hours 108 and 120 all three ensembles were more accurate than the GFS deterministic model.

For this route the aircraft crossed the polar front jet stream twice, in addition to having a majority of the flight traveling over areas of the world with very few

surface-based observations. This afforded the opportunity for considerable uncertainty in the model's first guess analysis. This, coupled with the fact that this route was the longest of the five routes tested may explain why the RMSE values are among the highest of all aircraft and route combinations tested.

4.1.2.5 KTCM-RJTY.

Results for the second longest flight tested, Travis to Yokota, are discussed in this section. The ensembles' percent improvement over the deterministic model were the most consistent throughout all forecast hours compared to previous routes and aircraft combinations. All three ensembles out-performed the GFS deterministic model for the majority of the forecast hours. The meteorological features encountered during this flight are similar to those for the ETAR-KDOV route. The aircraft encountered the polar front jet stream at least once along the path. Additionally, there are few surface observations located in the Pacific Ocean. Both of these factors will add to the initial condition errors for the deterministic model, as well as the ensembles.

The route's best case, the C-5, is depicted in Figure 21. During forecast hours 12 to 36, the GEFS and GEPS were more accurate than the GFS deterministic model with RMSE improvements ranging from 7% to 30%. During this same time period, the MEPS remained more accurate with the exception of f36, where it equaled the GFS deterministic model. The GEPS was the best performing ensemble, as it remained more accurate than the GFS deterministic model for all forecast hours. The GEFS was less accurate than the GFS deterministic model for only one out of ten forecast hours, f96, where it was only 1% less accurate. The GEFS had the lowest percentages of improvement, ranging from 1% to 11%. The MEPS had the highest range of improvements, ranging from 3% to 25%.

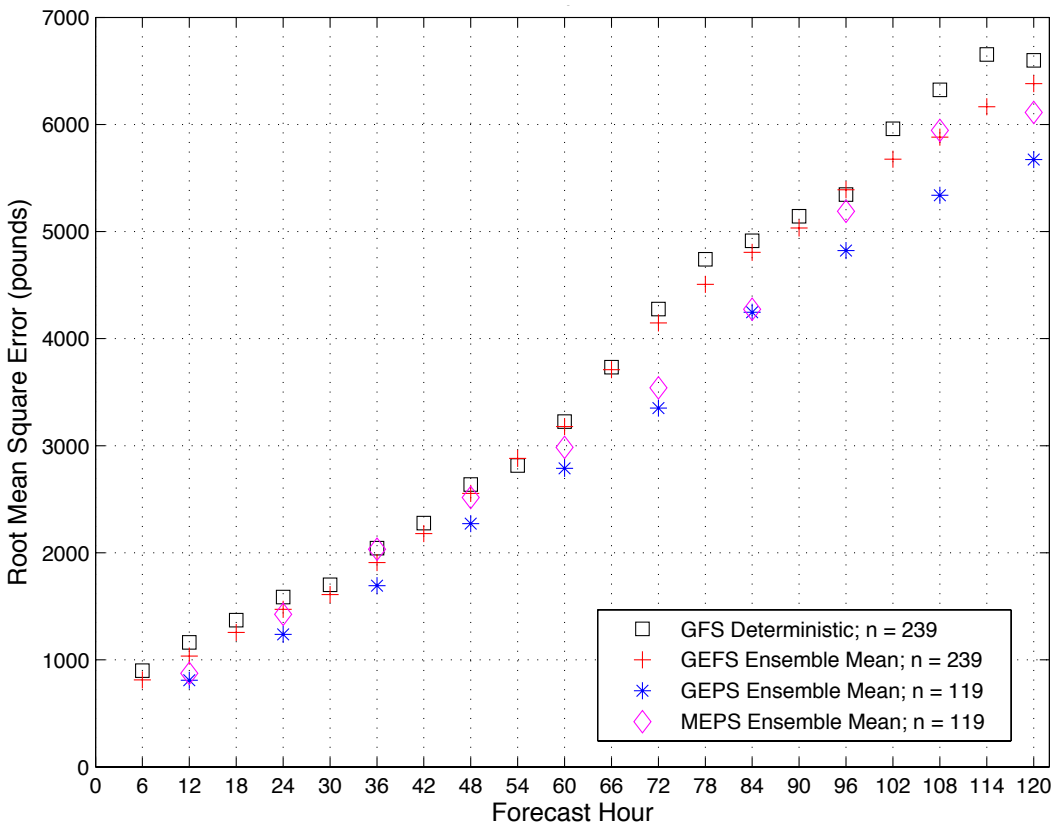

Figure 21. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 250mb (FL340) and flight route KTCM-RJTY. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the best case for this flight route. Best case means the ensembles showed the most improvement over the deterministic model.

Results for the worst case, the KC-135, are shown in Figure 22. The percent of improvement of the ensembles over the GFS deterministic model were similar to that of the best case (Figure 21) for all forecast hours. The GEPS was again the best performing ensemble overall, remaining more accurate than the GFS deterministic model for all forecast hours with improvements ranging from 8% to 31%. The GEFS and MEPS both out-performed the GFS deterministic for all forecast hours except f96 for GEFS and f36 for MEPS.

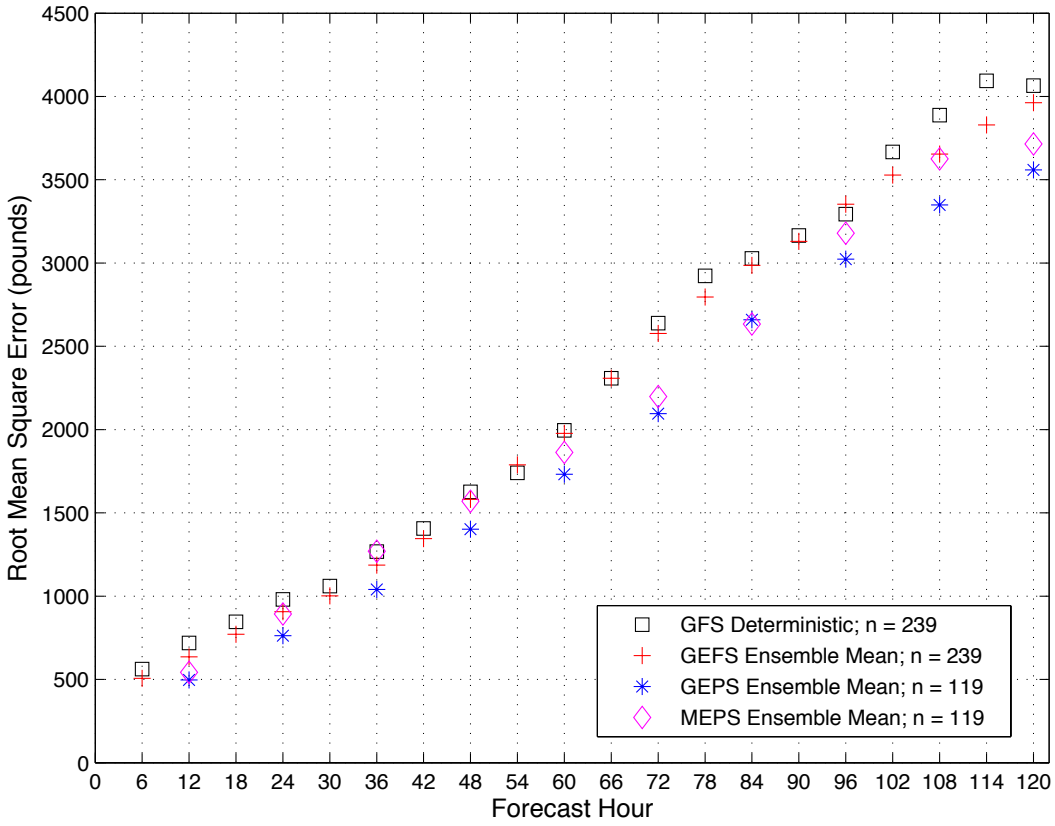

Figure 22. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-135 at cruise level 250mb (FL340) and flight route KTCM-RJTY. The timeframe of the data is 19 Sep – 17 Nov 2013. This represents the worst case for this flight route. Worst case means the ensembles showed the least improvement over the deterministic model.

4.1.2.6 FL300 (300mb).

At FL300 (300mb) there is only one ensemble, MEPS, available for comparison with the GFS deterministic model due to the non-availability of the 300mb level in the GEFS and GEPS output. For all four aircraft types on the McChord to Yokota and Travis to Hickam routes, the MEPS out-performed the GFS deterministic model for at least eight of the ten forecast hours. Figure 23 depicts the results for the KTCM-RJTY KC-10. The only forecast hour at which the GFS deterministic model is more accurate than the MEPS is f36. This is true also for the C-17 and C-5, while for the KC-135 the GFS was more accurate than the MEPS at

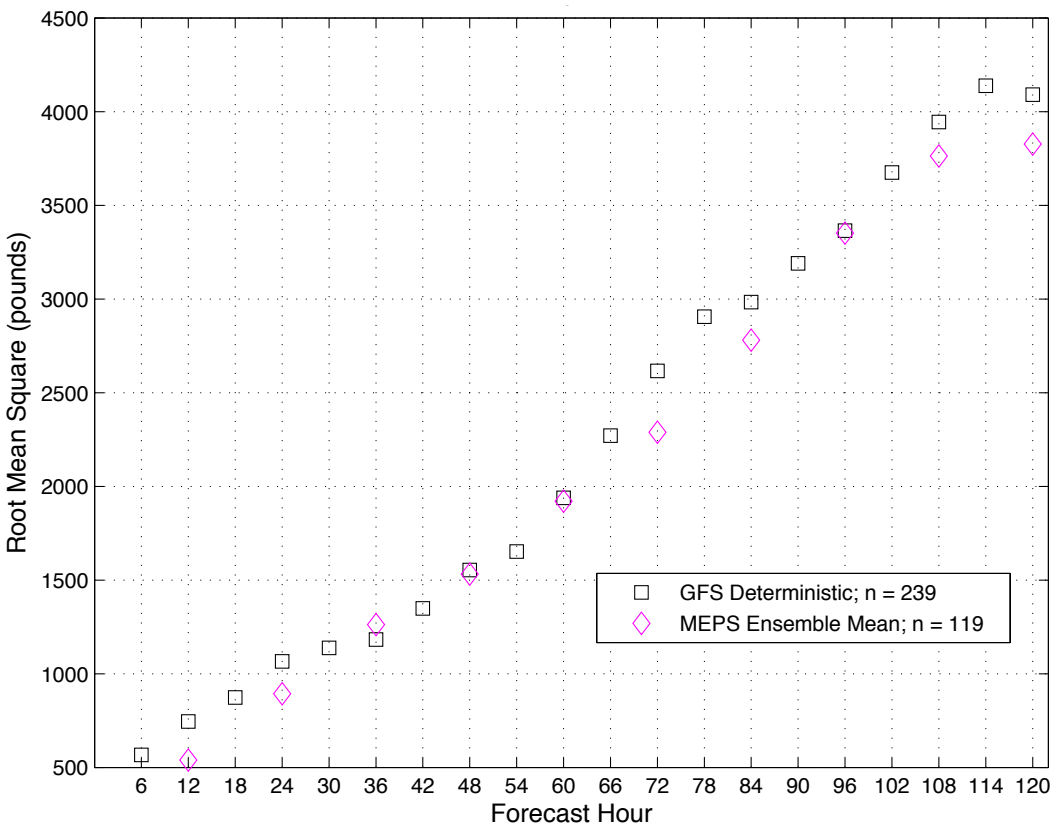

Figure 23. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a KC-10 at cruise level 300mb (FL300) and flight route KTCM-RJTY. The timeframe of the data is 19 Sep – 17 Nov 2013.

three forecast hours, f36 to f60. For the KSUU-PHIK route, all four aircraft only have one forecast hour, f84, at which the GFS out-performed the MEPS. For the remaining three flight routes, ETAR-KDOV, KCHS-KSUU, and KSUU-UCFM, the GFS generally was more accurate than the MEPS. Figure 24, which shows the KSUU-UCFM flight route for a C-5, is representative of the results for these three routes. For the KSUU-UCFM and KCHS-KSUU routes, the GFS deterministic model was more accurate than the MEPS for eight of the ten forecast hours. The MEPS was more accurate at the later forecast hours (f84 to f120). For the ETAR-KDOV route, the MEPS was more accurate than the GFS deterministic model only during the early forecast hours (f12 to f36).

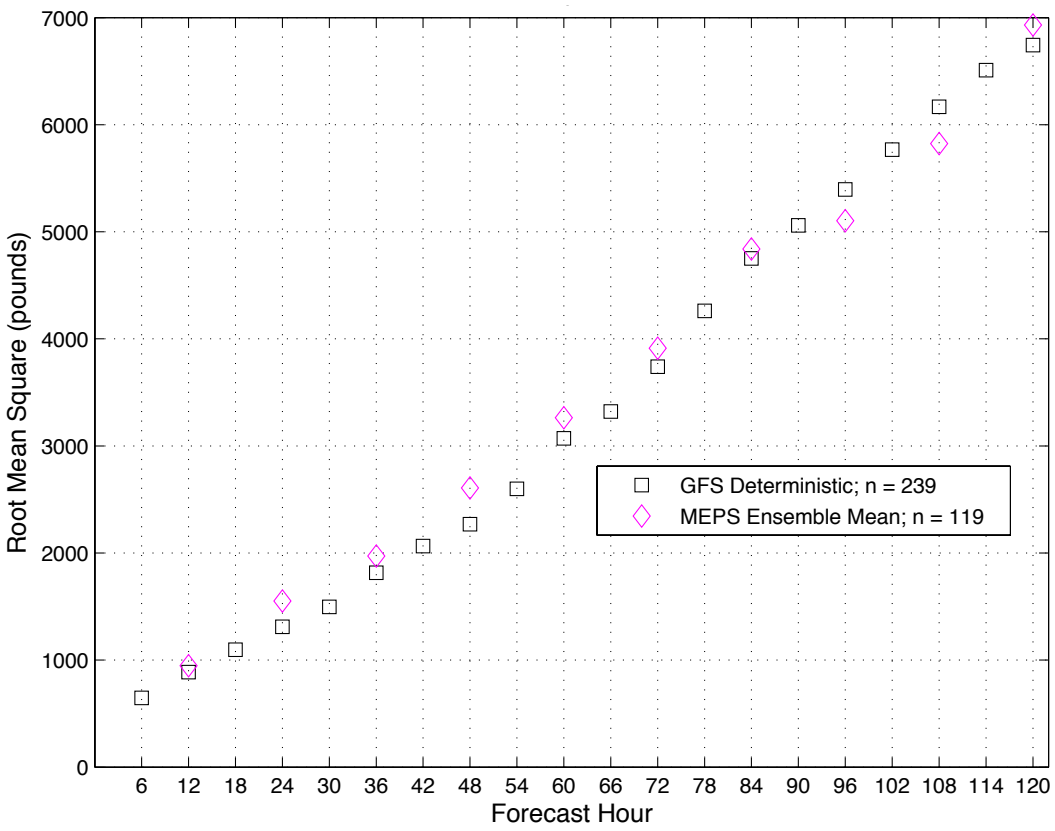

Figure 24. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-5 at cruise level 300mb (FL300) and flight route KSUU-UCFM. The timeframe of the data is 19 Sep – 17 Nov 2013.

4.1.2.7 FL240 (400mb).

Figure 25 depicts results for the KTCM-RJTY flight route for a C-17 at FL240 (400mb). This is the only route for which the MEPS was more accurate than the GFS deterministic model at all forecast hours and for all aircraft. Results for the ETAR-KDOV and KSUU-PHIK routes (not shown) were ambiguous. For the ETAR-KDOV route, the MEPS was more accurate than the GFS deterministic model during the early forecast hours (f12 to f36) but the GFS deterministic model became more accurate for the remainder of the forecast hours. KSUU-PHIK results were opposite, with the GFS deterministic model more accurate from f12 to f48 and the MEPS more accurate from f60 to f120. Figure 26 represents the results for the

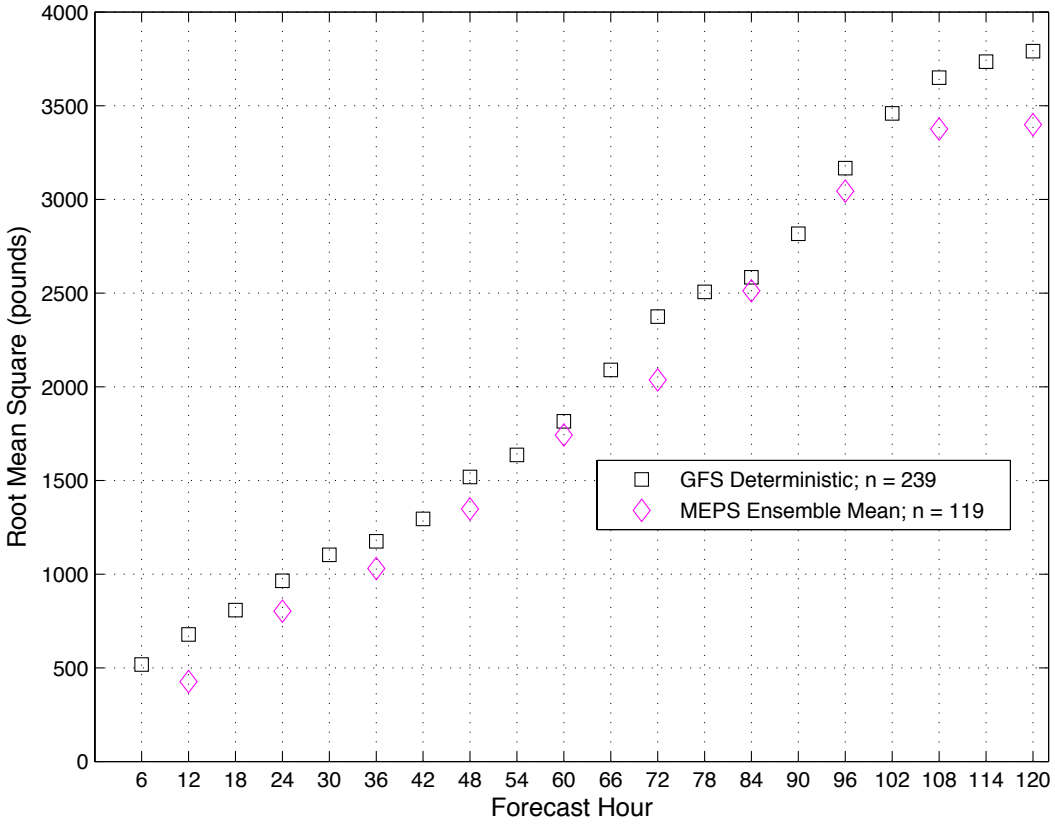

Figure 25. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-17 at cruise level 400mb (FL240) and flight route KTCM-RJTY. The timeframe of the data is 19 Sep – 17 Nov 2013.

KCHS-KSUU route. The KSUU-UCFM results were similar. On these two routes the GFS deterministic model out-performed the MEPS for nine of the ten forecast hours.

4.1.2.8 36-hr Forecast RMSE Comparison.

Table 3 contains the 36-hr forecast RMS fuel burn error values for the three ensembles and the GFS deterministic model. For each aircraft, cruise level, and route of flight the four numerical values represent the RMSE for each respective prediction system. The bold face values denote the prediction system with the highest accuracy (lowest RMSE) for that specific route, aircraft, and cruise level. The dashes denote unavailable data. Since the 300mb and 400mb pressure levels were not avail-

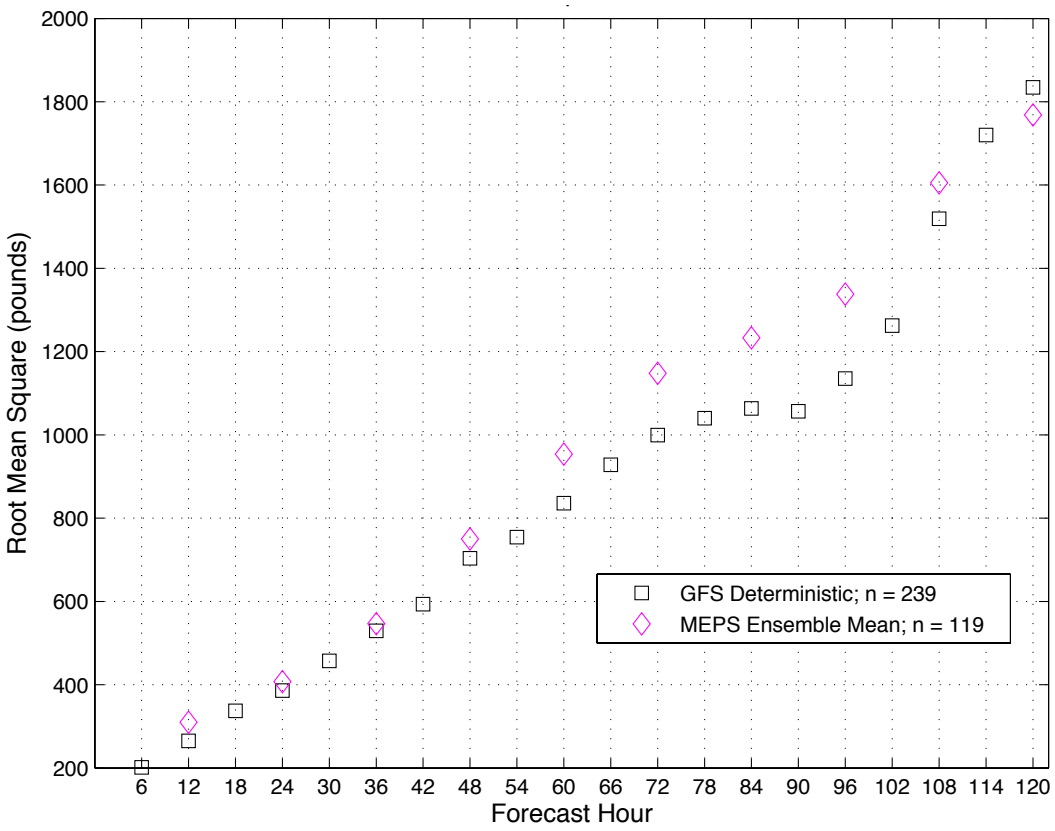

Figure 26. Root-mean square of fuel burn error in pounds as a function of model forecast hour for a C-17 at cruise level 400mb (FL240) and flight route KCHS-KSUU. The timeframe of the data is 19 Sep – 17 Nov 2013.

able from the GEFS and GEPS, the C-17, C-5, KC-10, and KC-135 do not have results for those two ensembles for routes flown at those cruise altitudes. Also, to remain consistent with the previous results, C-130 RMSE values were omitted for the KSUU-UCFM and KTCM-RJTY flight routes. Forecast hour 36 was chosen because it is the latest forecast hour typically used for flight planning. AMC typically produces ACFP flight plans/fuel computations between 6 and 12 hours prior to departure. Flight time for the AMC missions used in this study ranged from approximately 4 to 18 hours. The GFS and GEFS are available once every 6 hours, while the GEPS and MEPS are available once every 12 hours. Adding the flight plan lead time, flight duration, and model age produces a range of forecast hours between 16 and 42. With the GEPS and MEPS availability limited to every 12 hours the best

choice for comparison was the 36-hr forecast time, since it is close to f42 and includes the performance of the GFS deterministic model and all three ensembles. Although AMC typically only uses forecast hours 12 to 42 for flight plans it is important to analyze performance out to forecast hour 120 in order to quantify ensemble performance compared to the deterministic model for the occasions when a longer duration forecast is required. Appendix A contains similar RMSE data tables for forecast hours 12, 24, 60, 84, and 120.

Table 3. 36-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	287	**226**	—	—	523
	GEFS	279	240	—	—	465
	GEPS	**242**	230	—	—	**394**
	MEPS	289	228	—	—	516
C-17 250mb	GFS	818	720	**1286**	1528	1474
	GEFS	761	709	1414	1432	1202
	GEPS	**688**	702	1387	**1263**	**1121**
	MEPS	764	**681**	1419	1524	1144
C-17 300mb	GFS	**725**	**642**	**1355**	**1361**	1184
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	730	625	1468	1461	**1079**
C-17 400mb	GFS	**529**	**481**	**1052**	1175	962
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	546	505	1181	**1030**	**878**
C-5 250mb	GFS	1094	968	**1724**	2042	1961
	GEFS	1018	952	1892	1909	1608
	GEPS	**920**	943	1853	**1693**	**1506**
	MEPS	1021	**915**	1902	2033	1530

Table 3 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-5 300mb	GFS	**973**	862	**1815**	**1822**	1578
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	980	**839**	1971	1955	**1442**
C-5 400mb	GFS	**712**	**643**	**1412**	1572	1285
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	734	678	1586	**1379**	**1174**
KC-10 250mb	GFS	710	633	**1123**	1320	1264
	GEFS	663	624	1229	1235	1043
	GEPS	**598**	618	1203	**1093**	**979**
	MEPS	665	**599**	1244	1313	988
KC-10 300mb	GFS	**633**	567	**1187**	**1182**	1025
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	640	**551**	1291	1263	**936**
KC-10 400mb	GFS	**466**	**424**	**930**	1023	837
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	482	446	1044	**897**	**764**
KC-135 250mb	GFS	669	565	**1034**	1268	1236
	GEFS	616	555	1147	1186	975
	GEPS	**555**	552	1123	**1040**	**906**
	MEPS	615	**529**	1113	1269	942
KC-135 300mb	GFS	**575**	499	**1073**	**1115**	971
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	578	**483**	1145	1206	**884**

Table 3 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
KC-135 400mb	GFS	**415**	**370**	**815**	935	775
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	425	386	911	**828**	**704**

Table 3 indicates that, at f36, the ensembles out-performed the deterministic model in 87% of the cases. For the C-130 at 500mb (FL180) and the C-17, C-5, KC-10, and KC-135 at 250mb (FL340), the GEPS consistently performed best. The GEPS has nearly three times as many ensemble members as the GEFS and more than six times as many as the MEPS. The GEPS is a multi-model ensemble while the GEFS and MEPS are both single-model ensembles. Stensrud et al. (1999) found that for short-range forecasting the inclusion of two different models in an ensemble increases the ensemble spread, which may improve confidence in the accuracy of the ensemble mean. Additionally, Mylne et al. (2002) found that medium-range ensemble forecast skill improved when using a multi-model ensemble. An exception occurs for the KSUU-PHIK route, where the MEPS consistently had a lower RMSE at f36 (Table 3) and also for forecast hours 12, 24, 60, 84, and 96 as well (Appendix A). Thus it appears the MEPS may perform best in the tropics. The fuel burn RMSE improvements by all three EPSs for all aircraft types and routes of flight during forecast hours 12, 24, and 36 averages 10.25%. Thus, this preliminary study indicates that by using an ensemble mean wind forecast, AMC could potentially reduce flight plan fuel forecast errors.

4.2 Correlation of Fuel Burn Forecast Error with Forecast Additive Wind Spread

In order to test whether the ensemble additive wind spread might predict fuel burn forecast error, the same 65 aircraft/flight route combinations were analyzed. Table 4 shows the 36-hr correlation coefficients between absolute value of forecast fuel burn error and additive wind spread for all three ensembles. As Table 4 shows, there is no clear correlation. Figures 27 and 28 depict the strongest and weakest correlated cases. The x-axis depicts the ensemble additive wind spread in ms^{-1} and the y-axis depicts the absolute value of the forecast fuel burn error in pounds of fuel. The red line represents the best fit. The maximum correlation coefficient observed was 0.56 at f12 for the KSUU-PHIK KC-135 MEPS at 400mb (Figure 27). The lowest correlation coefficient (0.00) for all cases occurred at f66 for the ETAR-KDOV C-17 GEFS at 250mb (Figure 28).

The method used to calculate the additive wind spread in this study sums the combined wind spread values from grid point to grid point along the route of flight. This may not be the most optimal way to quantify wind forecast variability. The true spread one would want to measure is the spread in the flight times for each ensemble forecast (Kuchera 2014). The spread is often highly correlated from grid point to grid point. In such a case, the spread may be indicating uncertainty in the position of a meteorological feature rather than random wind variability. An example would be a situation in which every member of an ensemble forecasts a similar intensity for the speed maximum along a jet stream. Each member forecasts a similar intensity for the speed max, but a slightly different position. A flight through the area has a 100% chance of encountering the speed maximum and would feel the fuel burn impacts identically for each ensemble member. However, the additive wind spread would be large due to the jet streak position differences among ensem-

ble members. The large additive wind spread would indicate a large expected fuel burn error, while in reality the ensemble mean fuel forecast ends up being accurate. This may explain why this study did not find a strong correlation between the fuel burn error and the ensemble additive wind spread.

Table 4. Correlation between 36-hr fuel burn error absolute value and additive wind spread for each aircraft and route combination. Dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UAFM	KTCM-RJTY	ETAR-KDOV
C-130	GEFS	0.02	0.26	—	—	0.25
	GEPS	-0.03	0.16	—	—	-0.17
	MEPS	0.08	0.39	—	—	0.23
C-17 250mb	GEFS	0.14	0.15	-0.01	-0.13	-0.04
	GEPS	0.17	0.05	-0.01	-0.04	-0.17
	MEPS	0.07	0.13	-0.07	-0.14	-0.05
C-17 300mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	0.04	0.21	-0.05	-0.05	0.15
C-17 400mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	-0.01	0.20	-0.02	0.10	0.33
C-5 250mb	GEFS	0.14	0.15	-0.02	-0.13	-0.04
	GEPS	0.17	0.04	-0.01	-0.04	-0.17
	MEPS	0.07	0.12	-0.07	-0.14	-0.05
C-5 300mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	0.03	0.21	-0.05	-0.05	0.15
C-5 400mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	-0.01	0.19	-0.02	0.11	0.35

Table 4 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UAFM	KTCM-RJTY	ETAR-KDOV
KC-10 250mb	GEFS	0.14	0.14	-0.01	-0.13	-0.05
	GEPS	0.17	0.04	-0.01	-0.04	-0.19
	MEPS	0.07	0.11	-0.07	-0.14	-0.06
KC-10 300mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	0.04	0.20	-0.05	-0.05	0.14
KC-10 400mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	-0.01	0.19	-0.02	0.10	0.34
KC-135 250mb	GEFS	0.15	0.18	-0.02	-0.13	-0.01
	GEPS	0.17	0.09	-0.02	-0.06	-0.12
	MEPS	0.07	0.15	-0.08	-0.14	-0.03
KC-135 300mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	0.03	0.24	-0.06	-0.06	0.18
KC-135 400mb	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	-0.02	0.21	-0.02	0.11	0.36

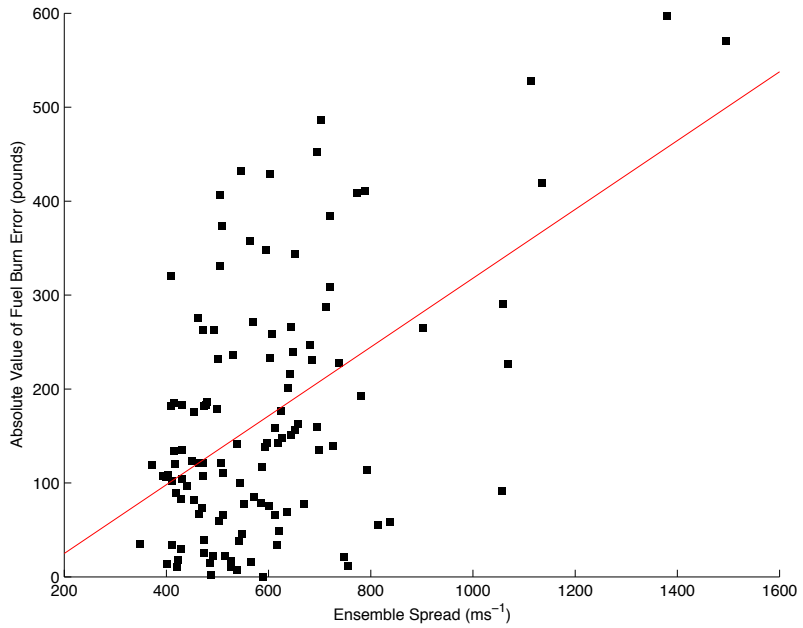

Figure 27. Absolute value of the fuel burn error in pounds versus the MEPS additive wind spread (ms^{-1}) at forecast hour 12 for a KC-135 at a cruise level of 400mb (FL240) and flight route KSUU-PHIK. The timeframe of the data is 19 Sep – 17 Nov 2013. This case had a correlation coefficient of .56 which is the best of all the cases tested.

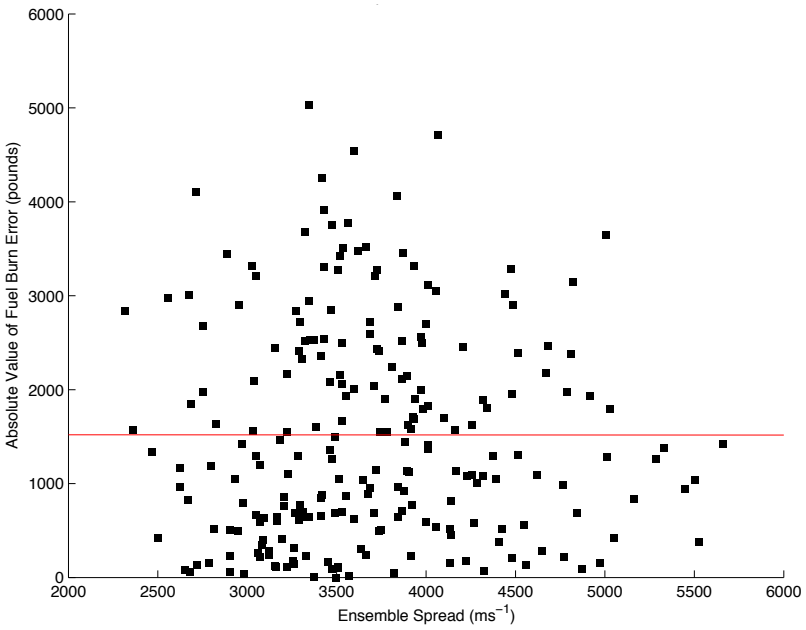

Figure 28. Absolute value of the fuel burn error in pounds versus the GEFS additive wind spread (ms^{-1}) at forecast hour 66 for a C-17 at cruise level 250mb (FL340) and flight route ETAR-KDOV. The timeframe of the data is 19 Sep – 17 Nov 2013. This case had a correlation coefficient of 0.00.

5. Conclusions

5.1 Summary

Over the past 60 years, major improvements and advances in NWP have made it the primary tool utilized by meteorologists today. Due to the chaotic nature of the atmosphere, weather models contain significant uncertainty as a result of analysis and model errors. Ensemble prediction systems enable meteorologists to quantify uncertainty in a forecast, so over the past decade meteorologists have gradually integrated EPSs into their forecasting and decision making processes. Numerous studies (Katz and Murphy 1997; Richardson 2000; Palmer 2002; Zhu et al. 2002) have shown the value of using probabilistic forecasts over deterministic or climatological information, especially for cost-loss decisions; AMC's airlift flight planning process, however, continues to use the GFS deterministic model as its primary forecast data source.

The two objectives of this research were to: (1) determine if an ensemble mean wind forecast is more accurate than a deterministic wind forecast for strategic airlift fuel planning, and (2) determine if a correlation exists between the ensemble additive wind spread and the error in calculated fuel burns. Currently the Automated Computer Flight Planning (ACFP) system used by AMC ingests GFS deterministic model weather data every 6 hours. These data (temperature, geopotential height, and the u and v-components of the wind), are then used to find the most optimum flight route. The objectives of this study were explored using an algorithm which calculated the amount of time and fuel required to fly a great circle route. As the aircraft traversed the route, its groundspeed, heading, and position were updated every minute using the u and v wind components from the nearest model grid point. Additionally, the combined ensemble wind spread was calculated

using the nearest grid point at each 1-minute time step and summed up along the route to give the additive wind spread.

The results indicate that ensemble mean wind forecasts generally provide more accurate fuel burn estimates than GFS deterministic model wind forecasts, producing an average improvement of 10.25% during forecast hours 12, 24, and 36. The increased accuracy of the wind forecasts may enable a reduction in the amount of reserve fuel AMC missions carry to accommodate for wind variability. With less reserve fuel required, the takeoff weight of the aircraft decreases, thus decreasing the amount of fuel burned throughout the flight. For the 23 different aircraft and flight route combinations tested at cruise levels 500mb (FL180) and 250mb (FL340) at least one of the three ensembles were more accurate than the GFS deterministic model in 96% of the cases during forecast hours 12 to 36. If the timeframe is extended to all forecast hours, the ensembles were more accurate than the GFS deterministic model in 89% of the cases. Among the individual ensembles, the GEPS consistently had the lowest RMSE (most accurate). At forecast hour 12, the GEPS was the most accurate model for all C-130 routes and four out of five of the C-17, C-5, KC-10, and KC-135 routes. During f24 and f36 the GEPS was always the most accurate ensemble for the C-130 routes and was the most accurate for at least three of the five routes for the remaining four aircraft. Since the GEPS (run twice daily) out-performs the GEFS (run four times daily) it can also be concluded from the limited cases in this study that a multi-model ensemble run at 12-hour intervals provides more accurate upper-level wind forecasts than a single-model ensemble run at 6-hour intervals. However, a study using a larger data set (spanning a full year) would provide more conclusive results. For the 40 different aircraft and flight route combinations tested at cruise levels 400mb (FL240) and 300mb (FL300), the GFS deterministic model generally out-performed the MEPS for at least three of the five

flight routes.

In regard to the second objective of this study, the results for the 65 different aircraft and flight route combinations tested showed that no correlation exists between the ensemble additive wind spread and the absolute value of the forecast fuel burn error. The highest correlation between the ensemble additive wind spread and the absolute value of the forecast fuel burn error found during this study was 0.56. The method used to calculate the total spread for each flight route may have contributed to the lack of correlation between the absolute value of the forecast fuel burn error and the ensemble additive wind spread.

5.2 Future Work

It is important to note that several simplifications were made during this study. To quantify potential cost savings, several improvements must be incorporated into this preliminary study: (1) expand the data set to span a full year; (2) incorporate 400mb (FL240) and 300mb (FL300) levels into standard GEPS output to enable comparison of all three EPSs for all flight levels of interest; (3) include temperature data along the flight route in order to more accurately calculate the true airspeed of the aircraft; (4) account for the variability of the weight of the aircraft throughout the flight; (5) incorporate the fuel burned during ascent and descent; and (6) develop a more robust method to calculate the ensemble forecast spread for each flight. Expanding the data set to at least a full year will not only yield a more statistically significant sample, but it will also afford the opportunity to investigate seasonal variability in deterministic model and ensemble performance.

Attempting to calculate a definitive cost savings based solely on the fuel burn RMSE reduction is very complex, especially since there are several other factors that affect the amount of fuel burned during a flight. However, the improvement in

the fuel burn forecast accuracy over the 60-day period of this study justifies further investigation of the value of implementing ensemble mean wind forecasts into the AMC flight planning process.

Appendix A. RMS Fuel Burn Error for Forecast Hours 12, 24, 60, 84, and 120

Table 5. 12-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	153	140	—	—	297
	GEFS	155	142	—	—	287
	GEPS	**103**	**103**	—	—	**190**
	MEPS	177	162	—	—	264
C-17 250mb	GFS	457	453	610	873	746
	GEFS	429	479	594	775	672
	GEPS	**334**	**392**	**478**	**602**	546
	MEPS	446	423	720	656	**540**
C-17 300mb	GFS	**370**	410	**661**	857	676
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	408	**397**	707	**623**	**452**
C-17 400mb	GFS	365	**264**	**478**	677	545
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**310**	281	531	**427**	**422**
C-5 250mb	GFS	613	605	818	1163	997
	GEFS	571	641	795	1036	904
	GEPS	**446**	**525**	**642**	**810**	733
	MEPS	598	568	963	875	**721**
C-5 300mb	GFS	**497**	551	**888**	1146	908
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	547	**533**	948	**834**	**604**

Table 5 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-5 400mb	GFS	**354**	**352**	**641**	901	732
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	415	377	708	**572**	**567**
KC-10 250mb	GFS	398	397	533	752	645
	GEFS	371	422	519	674	584
	GEPS	**291**	**345**	**421**	**524**	478
	MEPS	388	371	626	565	**467**
KC-10 300mb	GFS	**322**	360	**579**	745	589
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	358	**349**	617	**540**	**392**
KC-10 400mb	GFS	**233**	**233**	**423**	589	476
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	272	247	468	**372**	**367**
KC-135 250mb	GFS	369	352	491	719	614
	GEFS	345	375	479	636	549
	GEPS	**268**	**306**	**386**	**497**	446
	MEPS	358	331	585	542	**442**
KC-135 300mb	GFS	**294**	320	**524**	702	556
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	325	**310**	564	**504**	**368**
KC-135 400mb	GFS	**207**	**204**	**368**	539	443
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	243	216	407	**343**	**336**

Table 6. 24-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	210	184	—	—	419
	GEFS	207	181	—	—	401
	GEPS	**167**	**163**	—	—	**291**
	MEPS	204	207	—	—	380
C-17 250mb	GFS	639	591	929	1191	1199
	GEFS	600	548	949	1098	1024
	GEPS	**498**	543	**803**	**922**	866
	MEPS	598	**461**	1139	1068	**765**
C-17 300mb	GFS	549	513	**979**	1227	960
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**548**	**448**	1156	**1037**	**762**
C-17 400mb	GFS	**386**	**381**	**799**	965	796
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	407	406	881	**802**	**677**
C-5 250mb	GFS	854	792	1249	1587	1602
	GEFS	801	731	1270	1471	1370
	GEPS	**664**	729	**1077**	**1237**	1158
	MEPS	801	**620**	1527	1425	**1021**
C-5 300mb	GFS	738	686	**1311**	1642	1282
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**735**	**602**	1552	**1385**	**1022**

Table 6 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-5 400mb	GFS	**519**	**510**	**1071**	1290	1061
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	544	543	1184	**1074**	**907**
KC-10 250mb	GFS	555	517	809	1029	1031
	GEFS	521	480	828	953	889
	GEPS	**433**	475	**700**	**795**	751
	MEPS	521	**406**	995	918	**662**
KC-10 300mb	GFS	**479**	451	**856**	1066	831
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**479**	**394**	1014	**893**	**661**
KC-10 400mb	GFS	**339**	**334**	**703**	839	691
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	357	358	780	**697**	**590**
KC-135 250mb	GFS	515	466	747	981	1007
	GEFS	484	430	771	907	835
	GEPS	**400**	427	**651**	**763**	699
	MEPS	482	**358**	907	893	**633**
KC-135 300mb	GFS	**436**	400	**770**	1008	787
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**436**	**347**	905	**855**	**627**
KC-135 400mb	GFS	**302**	**294**	**615**	773	646
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	318	311	677	**644**	**548**

Table 7. 60-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	412	380	—	—	833
	GEFS	417	380	—	—	783
	GEPS	**406**	409	—	—	**683**
	MEPS	484	**358**	—	—	913
C-17 250mb	GFS	1042	1116	**2218**	2413	1798
	GEFS	**1013**	1102	2245	2383	**1626**
	GEPS	1070	1171	2358	**2086**	1732
	MEPS	1178	**964**	2279	2239	2083
C-17 300mb	GFS	**1017**	1066	**2297**	2243	**1890**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1165	**950**	2431	**2225**	2131
C17 400mb	GFS	**836**	825	**1743**	1815	**1583**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	954	**769**	1989	**1743**	1745
C5 250mb	GFS	1396	1502	**2975**	3225	2405
	GEFS	**1355**	1475	3005	3179	**2183**
	GEPS	1431	1572	3157	**2788**	2326
	MEPS	1576	**1292**	3060	2985	2788
C-5 300mb	GFS	**1364**	1428	**3071**	2996	**2528**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1563	**1277**	3263	**2973**	2849
C-5 400mb	GFS	**1122**	1108	**2340**	2430	**2123**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1280	**1036**	2670	**2334**	2335

Table 7 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
KC-10 250mb	GFS	907	980	**1931**	2084	1560
	GEFS	**882**	966	1958	2055	**1419**
	GEPS	932	1032	2058	**1799**	1513
	MEPS	1028	**846**	2004	1927	1805
KC-10 300mb	GFS	**890**	939	**2000**	1939	**1641**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1022	**838**	2138	**1921**	1848
KC-10 400mb	GFS	**735**	731	**1535**	1585	**1380**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	838	**681**	1754	**1520**	1520
KC-135 250mb	GFS	835	868	**1783**	1995	1474
	GEFS	**812**	861	1800	1977	**1322**
	GEPS	855	914	1890	**1723**	1395
	MEPS	939	**754**	1784	1863	1703
KC-135 300mb	GFS	**804**	821	**1820**	**1834**	**1550**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	921	**734**	1900	1836	1741
KC-135 400mb	GFS	**652**	630	**1354**	1447	**1270**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	746	**588**	1537	**1401**	1398

Table 8. 84-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	567	**522**	—	—	1207
	GEFS	**551**	529	—	—	1081
	GEPS	553	550	—	—	**1078**
	MEPS	650	524	—	—	1300
C-17 250mb	GFS	1440	1508	3460	3676	2849
	GEFS	**1391**	1551	3399	3601	2646
	GEPS	1529	1641	3418	**3181**	**2470**
	MEPS	1541	**1506**	**3372**	3202	3048
C-17 300mb	GFS	**1315**	**1335**	**3549**	3446	**3141**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1453	1381	3605	**3213**	3270
C-17 400mb	GFS	**1064**	1117	**2695**	2584	**2430**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1233	**1089**	3087	**2512**	2642
C-5 250mb	GFS	1929	2028	4639	4914	3809
	GEFS	**1863**	2080	4551	4807	3543
	GEPS	2048	2199	4578	**4246**	**3309**
	MEPS	2063	**2021**	**4529**	4274	4076
C-5 300mb	GFS	**1764**	**1795**	**4751**	4606	**4202**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1948	1852	4840	**4297**	4367
C-5 400mb	GFS	**1428**	1502	**3621**	3458	**3255**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1653	**1463**	4147	**3361**	3532

Table 8 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
KC-10 250mb	GFS	1253	1330	3013	3181	2474
	GEFS	**1213**	1361	**2965**	3105	2299
	GEPS	1334	1440	2987	**2735**	**2150**
	MEPS	1344	**1322**	2967	2767	2640
KC-10 300mb	GFS	**1152**	**1175**	**3102**	2983	**2720**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1271	1215	3174	**2781**	2827
KC-10 400mb	GFS	**935**	988	**2373**	2254	**2116**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1083	**962**	2722	**2188**	2296
KC-135 250mb	GFS	1152	**1177**	2789	3028	2318
	GEFS	**1113**	1214	2715	2986	2153
	GEPS	1221	1282	2722	2659	**2000**
	MEPS	1233	1180	**2624**	**2633**	2487
KC-135 300mb	GFS	**1042**	**1034**	2807	2819	**2575**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	1153	1075	**2805**	**2631**	2689
KC-135 400mb	GFS	**828**	852	**2091**	2065	**1946**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	964	**837**	2383	**2008**	2129

Table 9. 120-hr RMS Fuel Burn Error for each aircraft and route combination. The four rows correspond to the GFS, GEFS, GEPS, and MEPS respectively. The smallest RMSE for each aircraft/route combination is indicated in bold, while dashes denote non-availability of model output at that pressure level.

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
C-130	GFS	1006	847	—	—	2055
	GEFS	928	823	—	—	**1767**
	GEPS	886	838	—	—	1838
	MEPS	**882**	**765**	—	—	2124
C-17 250mb	GFS	2419	2858	4949	4939	4554
	GEFS	2426	2596	4867	4781	**4220**
	GEPS	2480	2500	**4811**	**4251**	4257
	MEPS	**2364**	**2311**	4941	4572	4978
C-17 300mb	GFS	2243	2412	**5028**	4722	**4923**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**2177**	**2078**	5129	**4407**	5211
C-17 400mb	GFS	1835	1772	**3930**	3791	**4107**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**1768**	**1569**	4103	**3399**	4225
C-5 250mb	GFS	3242	3840	6637	6599	6092
	GEFS	3247	3485	6518	6381	**5647**
	GEPS	3320	3353	**6447**	**5672**	5696
	MEPS	**3167**	**3100**	6634	6112	6554
C-5 300mb	GFS	3010	3244	**6743**	6315	**6585**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**2920**	**2787**	6931	**5896**	6960
C-5 400mb	GFS	2464	2384	**5281**	5082	**5504**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**2373**	**2109**	5508	**4544**	5648

Table 9 – continued from previous page

		KCHS-KSUU	KSUU-PHIK	KSUU-UCFM	KTCM-RJTY	ETAR-KDOV
KC-10 250mb	GFS	2113	2517	4340	4269	3958
	GEFS	2118	2282	4254	4121	**3665**
	GEPS	2163	2195	**4205**	**3650**	3702
	MEPS	**2066**	**2028**	4339	3967	4312
KC-10 300mb	GFS	1967	2128	**4414**	4091	**4268**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**1909**	**1827**	4507	**3827**	4502
KC-10 400mb	GFS	1613	1568	**3465**	3309	**3582**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**1555**	**1386**	3617	**2970**	3668
KC-135 250mb	GFS	1916	2218	3903	4065	3686
	GEFS	1932	2024	3854	3962	**3432**
	GEPS	1979	1957	**3835**	**3559**	3457
	MEPS	**1878**	**1810**	3874	3715	4084
KC-135 300mb	GFS	1764	1859	**3934**	3852	**4023**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**1716**	**1617**	4010	**3572**	4293
KC-135 400mb	GFS	1431	1354	**3034**	3027	**3297**
	GEFS	—	—	—	—	—
	GEPS	—	—	—	—	—
	MEPS	**1378**	**1206**	3167	**2703**	3416

References

Allen, M. S., 2009: Ambiguity in ensemble forecasting: Evolution, estimate validation and value. Dissertation, Naval Postgraduate School, Naval Postgraduate School, Monterey, CA 93943-5000.

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *Journal of Climate*, **9**, 1518–1530.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review*, **129**, 420–436.

Buizza, R., P. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Monthly Weather Review*, **133**, 1076–1097.

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **125**, 2887–2908.

Charney, J., 1955: The use of primitive equations of motion in numerical weather prediction. *Tellus*, **7**, 22–26.

Eckel, A., G. Cunningham, and D. Hetke, 2008: Exploiting forecast uncertainty for operational risk management. *Air and Space Power Journal*, **22 (1)**, 71–82.

Epstein, E., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739–759.

Hou, D., Z. Toth, W. Yang, and R. Wobus, 2011: A stochastic total tendency perturbation scheme representing model-related ucertainties in the NCEP global ensemble forecast system. *Tellus*.

Joint Working Group on Forecast Verification Research, 2012: Report to 5th joint scientific committee meeting of the world weather research programme. Tech. rep., World Meteorological Organization.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press.

Katz, R. and A. Murphy, 1997: *Economic value of weather and climate forecasts*. Cambridge University Press.

Keith, R. and S. M. Leyton, 2007: An experiment to measure the value of statistical probability forecasts for airports. *Weather and Forecasting*, **22**, 928–935.

Kuchera, E., 2012: Operational AFWA ensemble information. URL `https://weather.af.mil/confluence/display/AFPUBLIC/Operational+AFWA+Ensemble+Information`.

Kuchera, E., 2014: personal correspondence.

Leith, C., B. Adler, S. Fernbach, and M. Rotenberg, 1965: *Numerical simulation of the earth's atmosphere. Methods in Computational Physics*, Vol. 4. Academic Press.

Leith, C. E., 1974: Theoretical skill of monte carlo forecasts. *Monthly Weather Review*, **102 (6)**, 409–418.

Leutbecher, M. and T. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics*, **227**, 3515–3539.

Lewis, J. M., 2005: Roots of ensemble forecasting. *Monthly Weather Review*, **133**, 1865–1885.

Lorenz, E., 1965: On the possible reasons for long-period fluctuations of the general circulation. WMO-IUGG Symp. on Research and Development Aspects of Long-range Forecasting 66, World Meteorological Organization, WMO.

Lorenz, E., 1993: *The Essence of Chaos*. University of Washington Press.

McLay, J. and C. H. Bishop, 2010: A local formulation of the ensemble transform (ET) analysis perturbation scheme. *Weather and Forecasting*, **25**, 985–993.

Murphy, J. M., D. Sexton, D. Barnett, G. Jones, M. Webb, M. Collins, and D. Stainforth, 2004: Quantifying uncertainties in climate change using a large ensemble of global climate model predictions. *Nature*, **430**, 768–772.

Mylne, K., R. Evans, and R. Clark, 2002: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quarterly Journal of the Royal Meteorological Society*, **128**, 361–484.

Palmer, T. N., 2002: The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society*, **128 (581)**, 747–774.

Richardson, D., 2000: Skill and relative economic value of ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **126**, 649–668.

Richardson, L., 1965: *Weather Prediction by Numerical Process*. 1st ed., Dover.

Stensrud, D., H. Brooks, J. Du, M. Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Monthly Weather Review*, **127**, 433–446.

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at the NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, **74 (12)**, 2317–2330.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, **125**, 3297–3319.

Tracton, M. and E. Kalnay, 1993: Operational ensemble prediction at the national meteorological center: practical aspects. *Weather and Forecasting*, **8**, 379–398.

UCAR, 2004: Ensemble forecasting explained. URL `www.meted.ucar.edu/nwp/pcu1/ensemble/`.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79.

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society*, **83**, 73–83.

Vita

Captain Haley A. Homan was born in Creighton, NE, and graduated from Creighton Community High School in 2000. She attended Southeast Community College in Beatrice, NE from 2000 to 2002, earning National Junior College Athletic Association All-American honors in volleyball. In 2002, she transferred to the University of Oklahoma, where she earned a bachelor's degree in Meteorology and was commissioned as a Second Lieutenant via the Reserve Officer Training Corps in 2006. She was assigned to the 21st Operational Weather Squadron in Sembach, Germany from 2006 to 2008, rising to the position of C-Flight Production Supervisor. In the summer of 2008, she moved to Ellsworth Air Force Base, SD, where she served as Weather Flight Commander for the 28th Operations Support Squadron and was named the 28th Bomb Wing's Military Volunteer of the Year. From March 2011 to March 2012, Captain Homan commanded the 39th Operations Squadron Weather Flight at Incirlik Air Base, Turkey, where she was instrumental in safely establishing remotely-piloted aircraft operations. She was a member of the Air Force Women's Volleyball team in 2007, 2009, and 2010, as well as the Armed Forces Women's team in 2007 and 2009. Captain Homan entered the Applied Physics master's degree program at the Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, in the summer of 2012. Upon graduation, she will be assigned to the Air Force Weather Agency, Offutt Air Force Base, NE.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
27–03–2014	Master's Thesis	May 2012 — Mar 2014

4. TITLE AND SUBTITLE

Comparison of Ensemble Mean and Deterministic Forecasts for Long-Range Airlift Fuel Planning

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Homan, Haley A., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-14-M-15

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Mobility Command/POC: Frederick C. Wirsing HQ AMC/A3AW
402 Scott Dr Unit 3A1
Scott AFB, IL 62225-5302
DSN 779-3636, COMM 618-229-3636
Email: frederick.wirsing@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AMC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

This material is declared work of the U.S. Government and is not subject to copyright protection in the Untied States.

14. ABSTRACT

Implementing an ensemble mean forecast to aid in fuel planning for long-range strategic airlift has the potential to improve upon the deterministic forecasts currently used. More accurate wind forecasts could aid in a significant reduction in annual fuel costs for the DoD. This study focuses on the wind forecasts from the Global Forecast System (GFS) deterministic model and the ensemble mean wind forecasts from the Global Ensemble Forecast System (GEFS), Global Ensemble Prediction System (GEPS), and Mesoscale Ensemble Prediction System (MEPS) over a 60-day period from 19 Sep through 17 Nov 2013. The fuel burn and total spread was computed for five great circle flight routes and five aircraft using each model's wind data. The deterministic fuel burn error was then compared to the ensemble mean fuel burn error. For each of the flights investigated at cruise levels 500mb (FL180) and 250mb (FL340) the amount of reserve fuel required to account for the uncertainty in the wind forecasts was typically lower for the ensemble mean forecasts during forecast hours 12 to 48.

15. SUBJECT TERMS

Ensemble Mean Forecast, Deterministic Forecast, Ensemble Spread

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Robert S. Wacker, AFIT/ENP
U	U	U	U	85	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636, x4609; robert.wacker@afit.edu